

A stochastic framework of neurogenesis underlies the assembly of neocortical cytoarchitecture

Article (Accepted Version)

Llorca, Alfredo, Ciceri, Gabriele, Beattie, Robert, Wong, Fong Kuan, Diana, Giovanni, Serafeimidou-Pouliou, Eleni, Fernández-Otero, Marian, Streicher, Carmen, Arnold, Sebastian J, Meyer, Martin, Hippenmeyer, Simon, Maravall, Miguel and Marin, Oscar (2019) A stochastic framework of neurogenesis underlies the assembly of neocortical cytoarchitecture. *eLife*, 8. pp. 1-78. ISSN 2050-084X

This version is available from Sussex Research Online: <http://sro.sussex.ac.uk/id/eprint/88282/>

This document is made available in accordance with publisher policies and may differ from the published version or from the version of record. If you wish to cite this item you are advised to consult the publisher's version. Please see the URL above for details on accessing the published version.

Copyright and reuse:

Sussex Research Online is a digital repository of the research output of the University.

Copyright and all moral rights to the version of the paper presented here belong to the individual author(s) and/or other copyright owners. To the extent reasonable and practicable, the material made available in SRO has been checked for eligibility before being made available.

Copies of full text items generally can be reproduced, displayed or performed and given to third parties in any format or medium for personal research or study, educational, or not-for-profit purposes without prior permission or charge, provided that the authors, title and full bibliographic details are credited, a hyperlink and/or URL is given for the original metadata page and the content is not changed in any way.

A stochastic framework of neurogenesis underlies the assembly of neocortical cytoarchitecture

Alfredo Llorca^{1,2}, Gabriele Ciceri^{1,2,6}, Robert Beattie³, Fong Kuan Wong^{1,2}, Giovanni Diana^{1,2}, Eleni Serafeimidou-Pouliou^{11,2}, Marian Fernández-Otero^{1,2}, Carmen Streicher³, Sebastian J. Arnold⁴, Martin Meyer^{1,2}, Simon Hippenmeyer³, Miguel Maravall⁵ and Oscar Marín^{1,2*}

¹Centre for Developmental Neurobiology, Institute of Psychiatry, Psychology and Neuroscience, King's College London, London SE1 1UL, United Kingdom

²MRC Centre for Neurodevelopmental Disorders, King's College London, London SE1 1UL, United Kingdom

³Institute of Science and Technology Austria, Am Campus 1, 3400 Klosterneuburg, Austria

⁴Institute of Experimental and Clinical Pharmacology and Toxicology, Faculty of Medicine, University of Freiburg, Germany

⁵Sussex Neuroscience, School of Life Sciences, University of Sussex, Brighton BN1 9QG, United Kingdom

⁶Present address: Center for Stem Cell Biology, Memorial Sloan-Kettering Cancer Center, New York 10065, USA

*Correspondence should be addressed to O.M. (oscar.marin@kcl.ac.uk)

The cerebral cortex contains multiple areas with distinctive cytoarchitectonical patterns, but the cellular mechanisms underlying the emergence of this diversity remain unclear. Here, we have investigated the neuronal output of individual progenitor cells in the developing mouse neocortex using a combination of methods that together circumvent the biases and limitations of individual approaches. Our experimental results indicate that progenitor cells generate pyramidal cell lineages with a wide range of sizes and laminar configurations. Mathematical modelling indicates that these outcomes are compatible with a stochastic model of cortical neurogenesis in which progenitor cells undergo a series of probabilistic decisions that lead to the specification of very heterogeneous progenies. Our findings support a mechanism for cortical neurogenesis whose flexibility would make it capable to generate the diverse cytoarchitectures that characterize distinct neocortical areas.

The mammalian cerebral cortex contains a wide diversity of neuronal types heterogeneously distributed across layers and regions. The most abundant class of neurons in the cerebral cortex are excitatory projection neurons, also known as pyramidal cells (PCs). In the neocortex, PCs can be further classified into several subclasses with unique laminar distributions, projection patterns and electrophysiological properties (Greig et al., 2013; Jabaudon, 2017; Lodato and Arlotta, 2015), and currently available data suggest that several dozen distinct transcriptional signatures can be distinguished among them (Tasic et al., 2018). The relative abundance of the different types of PCs largely determines the distinct cytoarchitectonical patterns observed across different regions of the mammalian neocortex (Brodmann and Gary, 2006).

The diversity of excitatory neurons emerges from progenitor cells in the ventricular zone (VZ) of the developing neocortex known as radial glial cells (RGCs) (Malatesta et al., 2000; Miyata et al., 2001; Noctor et al., 2001). RGCs divide symmetrically to expand the progenitor pool during early stages of corticogenesis. Subsequently, they undergo asymmetric cell divisions to generate clones of PCs directly or indirectly via intermediate progenitor cells (IPCs) (Lui et al., 2011; Taverna et al., 2014). The characteristic vertical organization of migrating PCs in the developing neocortex led to the “radial unit hypothesis”, which postulates that PCs in a given radial column are clonally related (Rakic, 1988). However, the precise mechanisms through which RGCs generate diverse cytoarchitectonical patterns throughout the neocortex remain to be elucidated.

The most commonly accepted view of cortical neurogenesis is based on the notion that RGCs are multipotent and generate all types of excitatory neurons following an exquisite inside-out temporal sequence (Leone et al., 2008; Molyneaux et al., 2007; Rakic et al., 1994). Consistently, progenitor cells cultured in vitro reproduce the temporal

sequence of cortical neurogenesis (Gaspard et al., 2008; Shen et al., 2006), and genetic fate mapping experiments have shown that cortical progenitors identified by the expression of the transcription factors Fezf2 and Sox9 are multipotent in vivo (Guo et al., 2013; Kaplan et al., 2017). In contrast to this view, other studies have suggested the existence of fate-restricted cortical progenitors, which would only generate PCs for certain layers of the neocortex (Franco et al., 2012; García-Moreno and Molnar, 2015). However, the interpretation of these results remains a matter of controversy (Eckler et al., 2015; Gil-Sanz et al., 2015).

Our current framework for understanding cortical neurogenesis largely relies on studies that consider RGCs as a homogeneous population. Consistent with this view, recent clonal analyses of the developing neocortex led to the conclusion that progenitor cell behavior conforms to a deterministic program through which individual RGCs consistently generate the same neuronal output (Gao et al., 2014). This would suggest that variations in the organization of cortical areas would exclusively rely on mechanisms of lineage refinement at postmitotic stages, such as programmed cell death. Alternatively, the absence of detailed quantitated data of individual PC lineages or methodological caveats may have prevented the identification of a certain degree of heterogeneity in the neuronal output of individual RGCs.

In this study, we have used three complementary approaches to circumvent some of the intrinsic technical biases associated with each of the previously used methods to systematically investigate the clonal organization of PC lineages in the cerebral cortex. Our results provide a detailed quantitative assessment of the neurogenic fate of individual VZ progenitor cells that reveal a large diversity of PC lineage configurations. These findings support a stochastic model of cortical neurogenesis through which a limited number of

88 progenitor cell identities could generate the diverse of cytoarchitectonical patterns
89 observed in the neocortex.

Results

Retroviral tracing of pyramidal cell lineages

To study the cellular mechanisms underlying the generation of PCs in the neocortex, we analyzed the output and organization of neuronal lineages generated by individual progenitor cells. To this end, we first used replication-deficient retroviral vectors that integrate indiscriminately in mitotic cells but only identify cell lineages with fluorescent proteins following Cre-dependent recombination (Ciceri et al., 2013). To specifically label PC lineages, we injected a very low titer cocktail of conditional reporter retroviruses (*rv::dio-Gfp* and *rv::dio-mCherry*) into the lateral ventricle of *Neurod6*^{Cre/+} mouse embryos (also known as *Nex-Cre*), in which Cre expression is confined to postmitotic PCs (Goebbels et al., 2006) (*Figure 1a*). Using this approach, we achieved sparse labeling and avoided biasing the tagging of progenitor cells by the expression of specific genetic markers (Cepko et al., 2000).

To identify the developmental stage at which progenitor cells become neurogenic in the cortex, we injected retroviruses at different embryonic days (E9.5 to E14.5) and analyzed the organization of individual PC clusters at postnatal day (P) 21 (*Figure 1a*). Since a single copy of the viral vector is stably integrated into the host genome, retroviral infection leads to the labeling of only one of the two daughter cells resulting from the division of the infected progenitor cell. Consequently, infection of progenitor cells in the ventricular zone (VZ) of the pallium labels PC lineages in three main configurations depending of the mode of division of the infected progenitor (*Figure 1b*): (1) a large cluster containing more than one lineage, which results from the infection of a self-renewing progenitor cell dividing symmetrically; (2) a single lineage, which results from the infection of a progenitor cell undergoing its last symmetric division; and (3) a partial

lineage, which results from a neurogenic division of a progenitor cell. In this later case, partial lineages may contain the majority of neurons in the clone, if integration occurs in the progenitor cell, or one or two neurons, if the integration occurs in a neuron or an IPC, respectively.

We observed clusters of neurons with the characteristic morphology of PCs at all stages examined. Systematic mapping at P21 revealed very sparse labeling and widespread distribution of clones throughout the entire neocortex (*Figure 1c–c'* and *Figure 1—figure supplement 1*). The spatial segregation of the lineages was confirmed by the virtual absence of green and red clones within 500 μ m of each other in all experiments analyzed (*Figure 1—figure supplement 1*). We quantified the number of PCs per clone at P21 following viral infection at different embryonic stages and observed that lineages contain progressively smaller progenies (*Figure 1d*). This is consistent with the notion that VZ progenitors undergo proliferative symmetric cell divisions early during corticogenesis before they become neurogenic and begin self-renewing via asymmetric divisions (Gotz and Huttner, 2005; Kriegstein and Gotz, 2003). Since neurogenic divisions label one or two neurons in 50% of the cases (*Figure 1b*), the fraction of one- and two-cell clones found after retroviral infection is indicative of the proportion of neurogenic VZ progenitor cells at each embryonic stage. We observed that these clones represent ~50% of the lineages at E12.5 (*Figure 1e*). Consistent with previous reports using other methods (Gao et al., 2014), these results indicated that the onset of cortical neurogenesis begins immediately before E12.5, and that at this stage most VZ progenitor cells are already neurogenic. Thus, we focused subsequent analyses on this stage.

Cortical progenitors exhibit heterogeneous neurogenic output

We first examined lineages labeled at E12.5 that contained more than two cells, which correspond to the progeny of a VZ progenitor cell (*Figure 2a*). Consistent with classical models of cortical neurogenesis, we found that most VZ progenitor cells (63%) infected with retroviruses at E12.5 produce translaminal lineages containing neurons in both deep (V and VI) and superficial (II-III and IV) layers of the neocortex (*Figure 2b,e*). However, we also observed a substantial fraction of lineages in which PCs were confined to either deep (*Figure 2c,e*) or superficial (*Figure 2d,e*) layers (15% and 22%, respectively).

The distribution of single-cell and two-cell clones following infection of VZ progenitor cells at E12.5 further support the existence of cortical lineages restricted to superficial layers of the neocortex. As expected from the normal progression of neurogenesis in translaminal lineages, most single-cell and two-cell clones (which result from the labeling of a neuron or an IPC, respectively) were located in deep layers of the cortex (*Figure 2f-m*). However, in these experiments we also identified a small fraction of single-cell and two-cell clones in superficial layers of the neocortex (*Figure 2f-m*). This suggested that some VZ progenitor cells generate PCs for superficial layers of the neocortex in their earliest neurogenic divisions.

We noticed that the clonal size of laminar-restricted lineages is typically smaller than that of translaminal clones (*Figure 2e*). One explanation for this difference could be that laminar-restricted lineages represent sub-clones resulting from the labeling of IPCs that undergo more than one round of cell division, generating four to five neurons with a laminar-restricted distribution. To test this hypothesis, we carried out lineage tracing experiments at single cell resolution using low-dose tamoxifen administration in *Tbr2*^{CreERT2/+};RCL-Gfp pregnant mice at E12.5 (*Figure 2—figure supplement 1a*), which

led to the sparse labeling of IPCs and their progenies (Pimeisl et al., 2013). We analyzed 73 IPC-derived lineages at P21 and exclusively found one-cell and two-cell clones, with no evidence for larger clones within our sample (*Figure 2—figure supplement 1b–d*).

Although the existence of IPCs that undergo more than one cell division cannot be completely excluded, these results indicated that this is not common at this developmental stage. Consequently, IPCs are unlikely to be the origin of laminar-restricted lineages.

A small fraction of progenitors generates laminar-restricted lineages

Our retroviral tracing experiments suggested that the neurogenic output of neocortical VZ progenitor cells is significantly more heterogeneous than previously described, including translaminar, deep- and superficial-layer restricted lineages. However, several technical limitations may contribute to the observation of laminar-restricted lineages, as retroviral tracing may lead to the incomplete labeling of neuronal lineages. For example, the existence of deep layer-restricted lineages might be due to the silencing of the viral cassette after a few rounds of cell division (Cepko et al., 2000), which would prevent the expression of GFP or mCherry in superficial layer PCs. There are also alternative explanations for the observation of superficial layer-restricted lineages in the retroviral tracing experiments. First, infected progenitors might have become neurogenic at slightly earlier stages and have already produced a wave of deep layer PCs before infection, which would therefore not be labeled by the retrovirus. Second, the entire set of deep layer neurons might have been generated during the first neurogenic division of a VZ progenitor cell, which would not be labeled in some cases due to the retroviral integration mechanism.

To overcome these technical limitations, we took advantage of the Mosaic Analysis with Double Markers (MADM) technique, a genetic method widely used to fate-map cellular lineages at high resolution (Hippenmeyer et al., 2010; Zong et al., 2005). We used

the *Emx1-CreER^{T2}* mice (Kessaris et al., 2006) to induce MADM sparse labeling of VZ progenitor cells following tamoxifen administration at E12.5 (*Figure 3a*). We specifically focused our analysis on G2-X MADM segregation events that result in the labeling of an unbalanced number of daughter cells with either green or red fluorescent proteins and report the outcome of asymmetric divisions in VZ progenitor cells (Zong et al., 2005). Consistent with the retroviral lineage tracing experiments, we found that the vast majority of MADM lineages adopt a translaminar configuration (*Figure 3b,d*). In addition, we also identified some lineages in which PCs were confined to layers V and VI, thereby confirming the existence of cortical lineages restricted to deep layers of the neocortex (*Figure 3c,d*). We observed that the fraction of deep layer-restricted lineages labeled with MADM (~7%) is smaller than that obtained with retroviral tracing (15%), which suggested that reporter silencing might exist in some clones in the retroviral lineage tracing experiments. In contrast, we did not recover a significant number of superficial layer-restricted lineages in MADM experiments (*Figure 3d*).

The MADM experiments suggested that the observation of superficial layer-restricted lineages in retroviral experiments might be artifactual, a result of the incomplete retroviral labeling of neuronal lineages. We reasoned that if this were the case, the analysis of the MADM sub-clones (i.e., only one of the two colors in the lineage) containing more than two cells should lead to a similar fraction of ‘artifactual’ lineages, since these would essentially correspond to those labeled by retroviral infection missing the first division of VZ progenitor cells (*Figure 3—figure supplement 1*). This analysis indeed identified a small fraction of MADM sub-clones as ‘apparent’ superficial layer-restricted lineages (~12%), which was nevertheless significantly smaller than those identified in the retroviral

experiments (22%). This indicated that although some of the superficial layer-restricted lineages observed in retroviral experiments were artifactual, others might not be.

One important difference between both approaches is that MADM G2-X recombination events occur exclusively in mitotic cells (Zong et al., 2005), while retroviral labeling does not strictly depend on cell division. Retroviruses require cell division for their integration into the genome, but the infection is independent of cell cycle stage (Cepko et al., 2000). Thus, we hypothesized that MADM may not consistently label a fraction of quiescent or slowly dividing progenitors, which could otherwise be targeted by retroviral infection. To test this idea, we carried out a new set of lineage tracing experiments using a third, complementary method. In brief, we traced cortical lineages at single cell resolution using low-dose tamoxifen administration in *Emx1-CreER^{T2};RCL-Gfp* (*RCL-Gfp* also known as RCE) pregnant mice at E12.5 ([Figure 4a](#)), in which labeling of VZ progenitor cells should be independent of cell cycle dynamics. Since this method does not distinguish between lineages derived from symmetric or asymmetric cell divisions, we limited our analysis to lineages with a maximum of 12 cells, the larger clonal size of neurogenic lineages in the *Emx1-CreER^{T2};MADM^{TG/GT}* experiments (clones with more than 12 cells account for less than 5% of the neurogenic lineages and largely include [87%] the outcome of symmetrically-dividing progenitor cells).

Consistent with the other approaches, the majority of lineages (~75%) labeled by injection of *Emx1-CreER^{T2};RCL-Gfp* mice with low tamoxifen doses at E12.5 were translaminal ([Figure 4b,e](#)). We also confirmed that ~13% of the lineages were restricted to deep cortical layers ([Figure 4c,e](#)). In addition, we found that ~11% of the lineages consist of PCs confined to superficial layers of the neocortex ([Figure 4d,e](#) and [Figure 4—figure supplement 1](#)). In sum, the combined results of three different sets of lineage tracing

experiments suggested that translaminal (~80%), deep layer-restricted (~10%) and superficial layer-restricted (~10%) lineages are generated at the onset of neurogenesis in the developing neocortex.

Pyramidal cell lineages acquire diverse configurations

We next explored the precise organization of cortical lineages derived from VZ progenitor cells at E12.5. In lineage tracing experiments using *Emx1-CreER^{T2};RCL-Gfp* mice (**Figure 5a**), we observed that only about a quarter of traced lineages contains neurons in every cortical layer from II to VI, and every other clone lacks neurons in one or multiple cortical layers (**Figure 5b–d,f**). For instance, a significant proportion of translaminal lineages lack PCs in layer V (**Figure 5b,f**) or layer IV (**Figure 5c,f**) but, considered collectively, PC lineages adopt every possible configuration of laminar distributions in the neocortex (**Figure 5f**). The heterogeneous organization of cortical lineages was not exclusively observed in the experiments performed in *Emx1-CreER^{T2};RCL-Gfp* mice; similar results were obtained in retroviral lineage tracing (**Figure 5—figure supplement 1a–e**) and MADM experiments (**Figure 5—figure supplement 1f–j**). Although the clonal size of these lineages also exhibits great heterogeneity, it seems to follow a bimodal distribution, with maximums at approximately 4 and 8 cells (**Figure 5e**). Intriguingly, these two maximums largely correspond to restricted and translaminal lineages respectively, reinforcing the previously described link between clonal size and laminar configuration.

We characterized the organization of translaminal lineages with PCs in every layer by quantifying the relative proportion of neurons in deep and superficial layers. This analysis revealed that these cortical lineages typically showed a bias toward the production of PCs for superficial layers, although a minority of lineages displayed a preference toward deep layers or a balanced distribution across deep- and superficial layers (**Figure 5g**). In

general, the total amount of cells in superficial and deep cortical layers was slightly anti-correlated. To further explore the molecular diversity of PCs in these lineages, we stained P21 brain sections from *Emx1-CreER^{T2};RCL-Gfp* mice induced at E12.5 with antibodies against Ctip2 and Satb2, two transcription factors whose relative expression defines different types of PCs with unique patterns of axonal projections (Greig et al., 2013; Lodato and Arlotta, 2015). We identified four PC subclasses based on the expression of these markers and their laminar distribution (**Figure 5—figure supplement 2**): Cortico-cortical projection neurons (CCPN), subcerebral projection neurons (SCPN), cortico-thalamic projection neurons (CThPN) and heterogeneous projection neurons (HPN) (Harb et al., 2016). Using this classification, we found that nearly a quarter of all-layer translaminar lineages were composed exclusively by CCPNs, while multiple different combinations of PC identities comprise the remaining lineages (**Figure 5h**). Of note, only a minor fraction of all cortical lineages contains the entire complement of subtypes identified. Altogether, our experiments revealed that PC lineages exhibit a great degree of heterogeneity in the number and identities they comprise.

Heterogeneous lineage configurations arise directly from neurogenesis

The observed heterogeneity in cortical lineages likely emerges during neurogenesis. However, it is possible that selective cell death of specific PCs might contribute to the heterogeneous organization of cortical lineages. Recent studies have shown PCs undergo apoptosis during early postnatal stages (Blanquie et al., 2017; Wong et al., 2018). To explore the contribution of cell death to the heterogenous configuration of cortical lineages, we labeled clones by injecting a low dose of tamoxifen in *Emx1-CreER^{T2};RCL-Gfp* mice at E12.5 and analyzed their laminar organization at P2, prior to the period of PC death (Wong et al., 2018). We detected no significant differences in the average clonal size

or in the relative frequency of P2 translaminar and laminar-restricted lineages compared to P21 (*Figure 5—figure supplement 3a–d*). In addition, we observed that the diversity of lineage patterns was remarkably similar between P2 and P21 (*Figure 5—figure supplement 3e*). We also noticed a tendency (χ^2 test, $p = 0.099$) for the fraction of lineages with PCs in every layer to be larger and the frequency of lineages lacking PCs in layer 5 to be smaller at P2 compared to P21 (*Figure 5—figure supplement 3e*). These experiments suggested that although cell death may have a subtle impact in refining the final diversity of lineages and their relative proportion, such heterogeneity should arise directly during the process of cortical neurogenesis.

Laminar densities do not predict lineage structure

The variability in size and composition of PC lineages raises questions about the developmental mechanisms underlying their genesis. We first asked whether lineage structure is relevant for cortical cytoarchitectural development. It is formally possible that the diversity in laminar composition of cortical lineages is simply the consequence of a random process of PC generation in which the only boundary condition is the relative number of PCs that populate each layer of the cortex. To test this, we used the lineages mapped in the primary somatosensory cortex (S1) of *Emx1-CreER^{T2};RCL-Gfp* mice, which are meant to collectively generate a common pattern of laminar densities. We randomly permuted the PCs obtained from the different lineages while maintaining each neuron's laminar identity and the total number of cells in each lineage. If lineage structure were to exclusively affect the control of PC laminar fractions at population level, permuted lineages should be expected to match experimental data. As expected, the permutation process left unaltered the clonal size distribution and number of cells per layer observed in the experimental data (*Figure 6—figure supplement 1a,b*). However, it failed to replicate

the observed anti-correlation in neuron numbers between superficial and deep layers (*Figure 6—figure supplement 1c*). In addition, we observed that the laminar configuration of the permuted lineages differed from the experimental data (data not shown). These results indicated that the laminar distribution of neurons within each lineage arises specifically from an organized pattern of neurogenesis.

Stochastic models reproduce the diversity of progenitor outputs

One possibility to explain lineage diversity is the existence of multiple different VZ progenitor cell types with restricted potential to generate specific classes of PCs (Franco and Muller, 2013). However, the observed heterogeneity in lineage configurations may also arise from equipotent VZ progenitor cells that are subject to stochastic factors controlling their output, as proposed for the retina (He et al., 2012). To establish the feasibility of the latter scheme, we used a Bayesian approach to model the outcome of cortical progenitor cells following stochastic developmental programs (Diana, 2019). This method involved using a set of probabilistic rules for generating lineages, and subsequently inferring the number of rules required for the assignment of all lineages observed in the experimental data (see Methods for details). To avoid having to account for variability potentially attributable to differences in the distribution of lineages across areas of the neocortex, we only considered experimental data obtained from lineages mapped in S1 of *Emx1-CreER^{T2};RCL-Gfp* mice. We reasoned that some lineage configurations observed in our experiments (such those containing cells exclusively in deep cortical layers) could derive from an early interruption of the developing lineage, reflecting an early terminal division, or a progenitor cell undergoing cell death after a few rounds of division. Since the genesis of such lineages might therefore not arise from specific developmental programs, these configurations were also not taken into account for the inference of the stochastic

rules governing this process. The Bayesian inference approach revealed that models using one or two progenitor types are sufficient to produce a diversity of lineage compositions as found in our experimental data (*Figure 6a,e,f*). However, the approach failed to reproduce the anti-correlation in cell numbers between superficial and deep layers found experimentally (*Figure 6d*) and tended to underestimate the fractions of superficial and small lineages in our data set (*Figure 6c,e*). This suggests that while simple stochastic processes acting mostly on a single homogeneous population of VZ progenitor cells can originate a vast diversity of outcomes as observed in our experiments, some experimental observations may arise from additional developmental programs and from features characteristic of the sequential process of neurogenesis.

A small number of progenitor identities underlies lineage diversity

Having established that the stochastic behavior of a small number of progenitor types could in principle account for lineage diversity, we next explored specifically how diversity can result from the sequential dynamics of stochastic neuron generation. To this end we simulated cortical progenitor behavior using models based on four basic rules derived from experimental knowledge (*Figure 7a* and *Figure 7—figure supplement 1a*). First, in silico progenitors would generate neurons for different layers sequentially, following the observed inside-out pattern. Second, each in silico progenitor would have a set, randomly selected number of opportunities to generate neurons in each layer. Third, for any progenitor the decision to generate a neuron would be probabilistic, with cell generation probabilities varying by cortical layer but equal for all opportunities within the same layer. Thus, a progenitor type was defined by its specific combination of cell generation probabilities across layers. Fourth, to simulate the chances of premature terminal division and/or progenitor death, we introduced a probabilistic chance of lineage

interruption at each opportunity for cell generation. In silico lineages generated using this model were then compared with the experimental lineages mapped in the primary somatosensory cortex (S1) of *Emx1-CreER^{T2};RCL-Gfp* mice. We set cell generation probabilities in each layer to match the total laminar fractions of PCs (*Figure 7b* and *Figure 7—figure supplement 1b*), as well as the clonal size distribution (*Figure 7c* and *Figure 7—figure supplement 1c*) observed in those lineages.

In agreement with the Bayesian approach, we found that a stochastic model based on a single, equipotent VZ progenitor cell (Model 1), i.e., a single set of cell generation probabilities, was able to reproduce the majority of experimentally observed lineage features. Modeled lineages recapitulated the existence of restricted lineages as well as all observed laminar configurations of translaminar lineages (*Figure 7—figure supplement 1e–f*). However, this model generated lineages with an exaggerated anti-correlation in the number of cells in superficial versus deep layers (*Figure 7—figure supplement 1d*) and failed to reproduce the bimodal distribution of clonal sizes, underestimating the fraction of small lineages (those containing 3–4 cells). In addition, the fraction of lineages restricted to superficial layers, which largely contribute to the small lineage sizes, was also underestimated (*Figure 7—figure supplement 1c,e*). These results suggested that the fraction of small superficial lineages is unlikely to arise from a single stochastic program common to all cortical progenitors.

We then generated a second model with two different sets of cell generation probabilities, defining two different progenitor populations (Model 2). In this model, the majority of progenitors belonged to a population generating the larger lineages, while a second, smaller population generated small progenies biased towards superficial layer PC fates (*Figure 7a*). We found that this model faithfully reproduced all the experimental

features in our data: total laminar fractions (*Figure 7b*), bimodal distribution of clonal sizes (*Figure 7c*) and negative correlation in superficial versus deep layers (*Figure 7d*). In addition, the relative proportions of translaminar and laminar-restricted lineages were identical to those measured experimentally (*Figure 7e*). Finally, translaminar modeled lineages exhibited similar laminar configurations to the experimental lineages (*Figure 7f*). In sum, mathematical modelling suggests that a stochastic mechanism of cortical neurogenesis based on two independent progenitor cell populations best approximates the experimental data.

Finally, we explored whether the proposed stochastic model with two progenitor populations (Model 2) would be able to generate different ratios of layer-specific neurons under different circumstances (i.e., different cell generation probabilities), which would robustly account for the emergence of cytoarchitectural differences across neocortical areas. To this end, we quantified the fraction of PCs in each layer of the primary somatosensory (S1) and visual (V1) cortices in *Neurod6^{Cre/+};Fucci2* mice, in which all PCs in the neocortex are labeled with a nuclear fluorescent marker. As expected, we found important differences in laminar cytoarchitecture between both regions (*Figure 7g*). Remarkably, we found that subtle tuning of generation probabilities for both areas was sufficient to replicate the different laminar ratios in silico (*Figure 7h*). This suggests the stochastic mechanisms of neurogenesis described here would suffice to generate the diverse cytoarchitectonical patterns observed across neocortical areas.

Discussion

Our results indicate that the output of individual progenitor cells in the developing mouse neocortex is much more heterogeneous than previously anticipated. Progenitor cells most frequently generate PCs for both deep and superficial layers of the neocortex, as suggested by previous studies. However, a sizable fraction of those lineages lacks PCs in one or several layers. In addition, the heterogeneous output of cortical progenitor cells includes lineages in which PCs are restricted to either deep or superficial layers. Mathematical modeling suggests that this wide diversity of outputs is compatible with a stochastic model of cortical neurogenesis. Such model represents a robust and adaptable mechanism for the assembly of neocortical cytoarchitecture.

Methodological considerations

Understanding how individual lineages contribute to the production and organization of PCs is essential to articulate a coherent framework of cortical development. The analysis of the output of progenitor cells in the developing rodent cortex expands over three decades and has relied on four approaches: retroviral labeling (Luskin et al., 1988; Noctor et al., 2001; Noctor et al., 2004; Price and Thurlow, 1988; Reid et al., 1995; Walsh and Cepko, 1988, 1992), mouse chimeras (Tan et al., 1998), MADM (Beattie et al., 2017; Gao et al., 2014) and genetic fate-mapping (Eckler et al., 2015; Franco et al., 2012; García-Moreno and Molnar, 2015; Gil-Sanz et al., 2015; Guo et al., 2013; Kaplan et al., 2017). These studies often led to contradictory results, which has prevented the emergence of a consistent model. The prevalent view is that each progenitor cell in the developing pallium is multipotent and generates a cohort of PCs that populate all layers of the neocortex except layer I (Eckler et al., 2015; Gao et al., 2014; Guo et al., 2013; Kaplan et al., 2017),

as originally conceived in the radial unit hypothesis (Rakic, 1988). In contrast, some authors have suggested that many cortical progenitor cells are fate-restricted to generate PCs that exclusively occupy deep or superficial layers of the neocortex (Franco et al., 2012; Franco and Muller, 2013; García-Moreno and Molnar, 2015; Gil-Sanz et al., 2015). Here, we have used three different methods (retroviral labeling, MADM and genetic fate-mapping) to investigate the clonal production of cortical neurons by capitalizing on the synergy that emerges from the advantages of each individual approach. Our results indicate that this multi-modal approach is required to comprehensively capture the complex behavior of progenitor cells in the developing cortex.

Retroviral labeling has two important limitations: it only labels hemi-lineages and is prone to silencing, which may prevent the identification of the entire progeny of a progenitor cell (Cepko et al., 2000). Conversely, retroviral labeling targets progenitor cells indiscriminately and, consequently, is not biased towards a particular genetic fate (Cepko et al., 2000), as is the case for genetic strategies. MADM, on the other hand, has the enormous advantage of identifying both sister cells resulting from a cell division. However, G2-X MADM events require progenitor cells to undergo cell division at the time of induction because it directly relies on Cre-dependent inter-chromosomal mitotic recombination (Zong et al., 2005). Our results revealed that MADM does not reliably label a small fraction of progenitor cells present in the pallial VZ at E12.5 that gives rise to cohorts of PCs exclusively located in superficial layers of the neocortex. These lineages were however observed in both retroviral labeling experiments and in genetic tracing experiments using the same genetic driver (*Emx1-CreER^{T2}*) as in the MADM experiments, which strongly suggests that some *Emx1*+ progenitor cells producing exclusively superficial layer PCs in the developing cortex are not targeted by the MADM approach.

We hypothesize that these progenitors might be quiescent or slow-dividing progenitors at this stage and become more active at later stages of development. Finally, although the use of genetic fate-mapping strategies (e.g., *Emx1-CreER^{T2};RCL-Gfp*) is a powerful method to investigate cortical lineages, it has the important constraint of not being able to distinguish between symmetric proliferating and asymmetric neurogenic divisions. This hampers the analysis of clonal sizes, which can be otherwise accurately assessed with MADM except for the lineages that are not detected with this method.

Diversity of neocortical lineages

Previous clonal analyses based on MADM lineage tracing experiments led to the suggestion that individual progenitor cells in the pallial VZ produce a unitary output of approximately 8 excitatory neocortical neurons distributed throughout superficial and deep layers of the neocortex (Gao et al., 2014). However, those studies failed to identify lineages with restricted laminar patterns (either deep or superficial layer restricted clones). Consequently, they also underestimated the fraction of lineages with relatively small clonal size. In contrast, our analysis of neurogenic lineages revealed a bimodal distribution of clonal sizes with defined peaks centered at approximately 4 and 8 cells, which largely correspond to the contribution of laminar-restricted and translaminar lineages, respectively.

Previous studies have suggested that some neocortical progenitor cells generate laminar-restricted lineages of PCs (Franco et al., 2012; Gil-Sanz et al., 2015). In our experiments, approximately one in six cortical progenitor cells generate laminar-restricted lineages. The existence of lineages restricted to deep layers of the neocortex was observed with all three methods used in this study. Although some variation exists in the relative fraction of deep layer-restricted lineages observed with the different approaches, these differences lie within the expected experimental noise considering the relatively small

number of lineages that belong to this category. In addition, both retroviral labeling and genetic fate-mapping experiments identified a fraction of cortical progenitor cells that generate PCs that exclusively populate the superficial layers of the neocortex. It is conceivable that these lineages reflect the output of progenitor cells that had already produced an earlier cohort of deep layer neurons prior labelling. If this were the case, one should expect to observe similar results in MADM experiments. In such experiments, however, we did not recover a significant fraction of superficial lineages. Therefore, the discrepancy between the results of genetic fate-mapping and MADM experiments, in which the same mouse strain is used as the driver for recombination (*Emx1-CreER^{T2}*), suggests that these fate-restricted lineages arise from progenitor cells that are not actively dividing at E12.5. This hypothesis is consistent with the identification of a population of self-renewing progenitors with limited neurogenic potential during the earliest phases of corticogenesis (García-Moreno and Molnar, 2015). The existence of superficial layer-restricted cortical lineages is further supported by the identification of IPCs as early as E12.5 that generate superficial layer PCs (this study and (Mihalas et al., 2016)), when the majority of deep layer PCs are being generated. Since IPCs derive from VZ progenitor cells (Haubensak et al., 2004; Miyata et al., 2004; Noctor et al., 2004) and cortical neurogenesis begins at these stages in the mouse, this observation reinforces the idea that some progenitors are tuned to generate superficial layer PCs from early stages of corticogenesis. Although it is formally possible that some translaminal lineages in the retroviral and genetic fate-mapping experiments could arise from symmetric divisions generating two laminar-restricted lineages (one superficial and one deep), this possibility is unlikely considering the results of the MADM experiments. The results of those experiments indicate that symmetric divisions at E12.5 generate at least one translaminal

sub-lineage in virtually all cases (58/59), suggesting that most translaminal lineages arise from progressive divisions of individual progenitor cells.

Our study also revealed that, independently of the laminar distribution, individual cortical progenitor cells generate lineages with very diverse combinations of PC types. Cortical progenitors are thought to undergo progressive changes in their competency to generate different layer-specific types of PCs (Desai and McConnell, 2000; Oberst et al., 2019; Rakic, 1974). Consistent with this idea, our results reveal that most cortical progenitors generate diverse types of excitatory neurons. However, since many cortical progenitor cells fail to generate neurons for at least one layer of the neocortex, the majority of cortical lineages does not include the entire diversity of excitatory neurons. In other words, the fraction of individual cortical lineages that would be considered as “canonical” – i.e., containing all three main classes of excitatory projection neurons (CCPN, SCPN and CThPN) – is significantly smaller than previously anticipated. Considering the variance in clonal size and lineage composition of neocortical lineages, our results indicate that cortical progenitor cells exhibit very heterogeneous patterns of neuronal generation and specification. This interpretation challenges the view that the neuronal output of RGCs is deterministic (Gao et al., 2014).

A stochastic model of cortical neurogenesis

Our results indicate that stochastic developmental programs, in which cortical progenitors undergo a series of probability-based decisions for the generation of the different PC fates, are capable of generating the wide diversity of lineage configurations observed in our experiments. Therefore, in spite of the great diversity of configurations that exist among individual neocortical lineages, our results suggest that their genesis does not require a corresponding heterogeneity in VZ progenitors. The model proposed here is somewhat

reminiscent of that described for the developing rodent and zebrafish retina (Gomes et al., 2011; He et al., 2012). In line with our findings, stochastic mechanisms based on a single set of probability rules explain the genesis of most, but not all neuronal types in the mammalian retina (Gomes et al., 2011).

It is presently unclear whether laminar-restricted lineages arise from a pool of progenitor cells separate from those generating translaminar lineages or should simply be considered as extreme examples of the enormous diversity of lineage configurations uncovered by our study. The generation of lineages restricted to deep layers might be due to premature terminal division or death of the progenitor cell (Blaschke et al., 1996; Mihalas and Hevner, 2018), as considered in our models, but the existence of superficial layer-restricted lineages is more difficult to explain. Moreover, our mathematical model best reproduces the complex cytoarchitecture of the neocortex when two distinct progenitor cell identities are considered. Previous studies have identified morphological heterogeneity among pallial VZ progenitor cells (Gal et al., 2006). However, there is limited evidence for important molecular differences among these cells (Mizutani et al., 2007; Pollen et al., 2014; Telley et al., 2016). In the absence of a definitive molecular signature, our results suggest that while a homogeneous population of progenitor cells following a common developmental program explains most of the observed outcomes, it fails to generate the fraction of small superficial lineages observed in the experiments. The introduction of a second population of progenitor cells is required to reproduce these lineages. Although these findings might suggest the existence of a small fraction of cortical progenitors tuned to preferentially generate superficial lineages, it should not be taken as a definitive proof of fate-restriction in cortical progenitors. In our model, replicating the

experimental data does not require such progenitors to be restricted, but simply biased towards generating superficial fates.

Our study suggests that progenitor cells in different cortical areas are likely constrained by different probabilistic rules, which would contribute to the generation of the diverse cytoarchitectonic patterns found across the neocortex. Although the number of lineages recovered from each cortical region is insufficient to provide conclusive evidence for major regional differences, lineages located in different cortical areas seem to exhibit features that reflect the local cytoarchitecture. For instance, lineages lacking layer IV neurons were abundantly found in the retrosplenial cingulate and motor cortices, where this layer is remarkably small. How and when stochastic neurogenic decisions are made remains to be elucidated, but they likely depend on the influence of extrinsic and intrinsic signals on parameters such as cell cycle length, the asymmetric inheritance of cell components, the generation of dividing (IPCs) versus postmitotic progeny, and the membrane potential of progenitor cells (Haydar et al., 2003; Lange et al., 2009; Pilaz et al., 2009; Roccio et al., 2013; Vitali et al., 2018; Wang et al., 2009). Local signals in different neocortical areas would contribute to the tuning of progenitor cell behaviors to output different cytoarchitectures without the requirement of regional-specific progenitor populations. Consequently, this model allows great flexibility in the generation of heterogeneous cortical cytoarchitectures without the requirement of a large number of progenitor identities. The specification of a very small number of progenitor cells with competence to adapt their neurogenic behavior to different probabilistic rules based on their location within the neocortical neuroepithelium represents the most parsimonious and robust mechanism for the generation of cortical circuitry.

561 Materials and methods

562 Key resources table

Reagent type (species) or resource	Designation	Source or reference	Identifiers	Additional information
Genetic reagent (<i>M. musculus</i>)	<i>Neurod6</i> ^{Cre}	PMID: 17146780	RRID: MGI:4429427	Dr Klaus Nave (MPI-EM)
Genetic reagent (<i>M. musculus</i>)	<i>Emx1-Cre</i> ^{ERT2}	PMID: 16388308 Jackson laboratory	Stock: 027784 RRID: IMSR_JAX:027784	
Genetic reagent (<i>M. musculus</i>)	<i>Tbr2</i> ^{CreERT2}	PMID: 23897762	RRID: MGI:5499789	Dr Sebastian J Arnold (University of Freiburg)
Genetic reagent (<i>M. musculus</i>)	<i>RCL-Gfp</i> , RCE	PMID: 19363146 Jackson laboratory	Stock: 32037 RRID: MGI:4420759	
Genetic reagent (<i>M. musculus</i>)	MADM ^{TG}	PMID: 21092859 Jackson laboratory	Stock: 013751 RRID: IMSR_JAX:013751	
Genetic reagent (<i>M. musculus</i>)	MADM ^{GT}	PMID: 21092859 Jackson laboratory	Stock: 013749 RRID: IMSR_JAX:013749	
Genetic reagent (<i>M. musculus</i>)	<i>RCL-Fucci2a</i>	PMID: 25486356	RRID: IMSR_HAR:6899	Dr Richard Mort (Lancaster University)
Cell line (<i>Homo sapiens</i>)	HEK293FT	Invitrogen	Cat: R700-07 RRID: CVCL_6911	
Transfected construct (<i>M. musculus</i>)	<i>rv::dio-eGFP</i>	PMID: 23933753 Addgene	ID: 87662	

Transfected construct (<i>M. musculus</i>)	<i>rv::dio-mCherry</i>	PMID: 23933753 Addgene	ID: 87664	
Antibody	Anti-GFP (chicken polyclonal)	Aves Lab	Cat: GFP-1020 RRID: AB_10000240	IF(1:2000)
Antibody	Anti-DsRed (rabbit polyclonal)	Clontech	Cat: 632496 RRID: AB_10013483	IF(1:500)
Antibody	Anti-mCherry (goat polyclonal)	Antibodies online	Cat: ABIN1440057 RRID: AB_11208222	IF(1:500)
Antibody	Anti-Ctip2 (rat monoclonal)	Abcam	Cat: AB18465 RRID: AB_2064130	IF(1:500)
Antibody	Anti-Satb2 (mouse monoclonal)	Abcam	Cat: AB51502 RRID: AB_882455	IF(1:500)
Antibody	Anti-Satb2 (rabbit polyclonal)	Abcam	Cat: AB34735 RRID: AB_2301417	IF(1:1000)
Antibody	Anti-Tle4 (goat polyclonal)	Gift from Dr Stefano Stifani (McGill University)		IF(1:200)
Antibody	Anti-chicken IgY-Alexa fluor 488 (goat polyclonal)	TermoFisher	Cat: A-11039 RRID: AB_2534096	IF(1:400)
Antibody	Anti-mouse IgG-Alexa fluor 647 (goat polyclonal)	TermoFisher	Cat: A-21240 RRID: AB_2535809	IF(1:400)
Antibody	Anti-mouse IgG-biotinilated (Horse polyclonal)	Vector Labs	Cat: BA2000 RRID: AB_2313581	IF(1:400)

Antibody	Anti-rat IgG-Alexa fluor 555 (goat polyclonal)	TermoFisher	Cat: A-21434 RRID: AB_2535855	IF(1:400)
Antibody	Anti-goat IgG-Alexa fluor 555 (donkey polyclonal)	TermoFisher	Cat: A-21432 RRID: AB_2535853	IF(1:400)
Antibody	Anti-rabbit IgG-Alexa fluor 488 (donkey polyclonal)	TermoFisher	Cat: A-21432 RRID: AB_2535853	IF(1:400)
Chemical compound, drug	Tamoxifen	Sigma-Aldrich	Cat: 85256	
Software, algorithm	Prism 8	GraphPad	RRID: SCR_002798	
Software, algorithm	MATLAB	Mathworks	RRID: SCR_001622	
Software, algorithm	RStudio	RStudio	RRID: SCR_000432	
Software, algorithm	Imaris 8	Bitplane	RRID: SCR_007370	

563

564 Mice

565 The following transgenic mouse lines were used in this study: *Neurod6^{Cre}* (Goebbels et al.,
566 2006) RRID: MGI:4429427, *Emx1-Cre^{ERT2}* (Kessaris et al., 2006) RRID:
567 IMSR_JAX:027784, *Tbr2^{CreERT2}* (Pimeisl et al., 2013) RRID: MGI:5499789, *RCL-Gfp*
568 (Sousa et al., 2009) RRID: MGI:4420759, *MADM^{TG}* (JAX 013751) RRID:
569 IMSR_JAX:013751, *MADM^{GT}* (JAX 013749) RRID: IMSR_JAX:013749 and *RCL-*

Fucci2 RRID: IMSR_HAR:6899. The MADM^{TG} and MADM^{GT} alleles were generated by inserting *dT_{N-term}-Gfp_{C-term}* and *Gfp_{N-term}-tdT_{C-term}* sequences in the *Hipp11* locus (Hippenmeyer et al., 2010). *RCL-Fucci2* are reporter mice in which a fluorescent ubiquitination-based cell cycle indicator (Fucci) consisting of *mVenus-hGem* and *mCherry-hCdt1* sequences linked by a T2A linker were flanked by loxP sites and inserted into the ROSA26 locus (Mort et al., 2014).

All adult mice were housed in groups and kept on reverse light/dark cycle (12/12 h) regardless of genotypes. Only time-mated pregnant female mice that have undergone in utero surgeries were housed individually. Both male and female mice were used in all experiments. In utero experiments were performed at different developmental stages that range from E9.5 to E14.5. For histological analyses, mice ages range from P2 to P30. All procedures were approved by King's College London and IST Austria, and were performed under UK Home Office project licenses, and in accordance with Austrian Federal Ministry of Science and Research license, and European regulations (EU directive 86/609, EU decree 2001-486). The day of vaginal plug was considered as embryonic day (E) 0.5 and the day of birth as postnatal day (P) 0.

Retroviral infection for clonal labelling

Cre-dependent conditional retroviral stocks encoding EGFP and membrane-bound mCherry reporters (*rv::dio-eGfp* and *rv::dio-mCherry*) (Ciceri et al., 2013) were produced as previously described (Tashiro et al., 2006). In brief, Moloney murine leukemia viruses (MoMLV) were produced by transfecting HEK293T cells (RRID: CVCL_6911) with retroviral plasmids (*dio-eGfp* or *dio-mCherry*, *pCMV-Vsvg*, and *pCMV-GAG-pol*) using lipofectamine 2000. Forty-eight hours post transfection the supernatant was collected, concentrated and purified by two sequential rounds of ultracentrifugation. The viral pellet

was re-suspended in sterile PBS and stored in aliquots at -80°C. Viral stocks for *dio-eGfp* and *dio-mCherry* were produced in the same plates and mixed before concentration by ultracentrifugation.

For in utero injections, pregnant females were deeply anesthetized with isoflurane and the abdominal cavity was incised to expose uterus. Conditional retroviruses were injected at low-titer into the telencephalic ventricles of E9.5, E10.5, E11.5, E12.5 and E14.5 mouse embryos using an ultrasound-guided imaging system (Visualsonic) coupled with a nanoliter injector as previously described (Ciceri et al., 2013; Pla et al., 2006). Some experiments were performed using the *rv::dio-eGfp* exclusively. After the procedure, the uterine horns were placed back in the abdominal cavity and the wound was surgically sutured. The female was then placed in a 32°C recovering chamber for 30 mins post-surgery before returning to standard housing conditions.

Inducible genetic clonal labelling

Emx1-Cre^{ERT2};RCL-Gfp and *Tbr2^{CreERT2/+};RCL-Gfp* pregnant females received a single intraperitoneal injection of low dose (1 ng/kg) tamoxifen dissolved in corn oil at E12.5. MADM clones were generated as described previously (Beattie et al., 2017; Hippenmeyer et al., 2010). In brief, timed pregnant females were injected intraperitoneally with tamoxifen dissolved in corn oil at E12.5 at a dose of 2-3 mg/pregnant dam. Live embryos were recovered at E18–E19 through caesarean section, fostered, and raised for further analysis at P21.

Histology

Postnatal mice were perfused transcardially with 4% paraformaldehyde (PFA) in PBS and the dissected brains were fixed for 2 hours at 4°C in the same solution. Brains were serially

sectioned at 100 μ m on a vibratome (VT1000S, Leica) or on a freezing microtome (SM 2010R, Leica) and free-floating coronal sections were then subsequently processed for immunohistochemistry as previously described (Pla et al., 2006).

The following primary and secondary antibodies were used: chicken anti-GFP (1:2000 Aves lab cat. no. GFP-1020, RRID:AB_10000240), rabbit anti-DsRed (1:500 Clontech cat. no. 632496, RRID:AB_10013483), goat anti-mCherry (1:500 Antibodies-Online cat. no. ABIN1440057, RRID:AB_11208222), rat anti-Ctip2 (1:500 Abcam cat. no. Ab18465, RRID:AB_2064130), mouse anti-Sab2 (1:500 Abcam cat. no. Ab51502, RRID:AB_882455), rabbit anti-Sab2 (1:1000 Abcam cat. no. Ab34735, RRID:AB_2301417), goat anti-Tle4 (1:200 gift from Stefano Stifani), anti-chicken IgY (H+L) 488 (1:400 Molecular Probes cat. no. A-11039, RRID:AB_2534096), anti-mouse IgG1 647 (1:400 Molecular Probes cat. no. A-21240, RRID:AB_2535809), anti-mouse IgG (H+L) biotinylated (1:400 Vector laboratories cat. no. BA-2000, RRID:AB_2313581), anti-rat IgG (H+L) 555 (1:400 Molecular Probes A-21434, RRID:AB_2535855), anti-goat IgG (H+L) 555 (1:400 Molecular Probes cat. no. A-21432, RRID:AB_2535853), and anti-rabbit 488 (1:400 Molecular Probes cat. no. A-21206, RRID:AB_2535853).

Imaging

Images were acquired using fluorescence microscopes (DM5000B, CTR5000 and DMIRB from Leica or Apotome.2 from Zeiss) coupled to digital cameras (DC500 or DFC350FX, Leica; OrcaR2, Hamamatsu) with the appropriate emission filter sets or in inverted confocal microscopes (Leica TCS SP8 and Zeiss LSM800 Airyscan).

In silico modelling of cortical development

All modelling of progenitor behavior was performed using MATLAB (MathWorks; RRID:SCR_001622). To avoid overfitting variability that could correspond to differences in progenitor behavior across cortical areas, simulations were compared to the lineages observed in primary somatosensory cortex (S1), using the *Emx1-CreER^{T2};RCL-Gfp* experiments. The structural similarity of the results of a model with the experimental data was assessed based on three parameters: proportion of cells per layer, clonal size distribution and Spearman correlation (r) values for number of cells in upper versus lower layers. For each parameter, we computed a normalized z-score measure by taking the difference between the experimental value and the average value across simulation repeats, and then dividing by the standard deviation across simulation repeats. Thus, z-score values over 1 would reflect a distance between experimental and modelled data larger than the standard deviation between simulation repeats.

To generate randomly permuted cortical lineages, neurons observed in the *Emx1-CreER^{T2};RCL-Gfp* experiments were permuted among lineages while maintaining their laminar identities. This operation was repeated 1,000 times, providing average and standard deviation values that were then used to compare with the experimental results.

Bayesian inference of progenitor types

To perform statistical inference on the number of categories required to explain the distribution of lineages throughout cortical layers, we employed a statistical model where N observed lineages are grouped in K progenitor types. Each type $t = 1, \dots, K$ is associated to a vector of four probabilities $p_t = \{p_t^{(III/IV)}, p_t^{(IV)}, p_t^{(V)}, p_t^{(VI)}\}$ representing the probabilities of any progenitor in the class to generate neurons in each of the four layers. We assume that each observed lineage can be assigned to a unique progenitor type based on its

occupancy distribution. Progenitor types are associated with frequencies f_t , reflecting how likely a lineage is to belong to type t . The occupancy probabilities and the relative frequencies for each type as well as the number of types K required can be obtained using Bayesian inference according to the Bayes' theorem

$$P(t_{1:N}, p_{1:K}, f_{1:K} | S) = \frac{\overbrace{P(t_{1:N}, S | p_{1:K}, f_{1:K})}^{\text{likelihood}} \cdot \overbrace{P(p_{1:K}, f_{1:K})}^{\text{prior}}}{\underbrace{P(S)}_{\text{marginallikelihood}}}$$

where S is the count matrix whose elements S_{ij} indicate how many neurons in lineage i belong to layer j . The Bayes' theorem provides the posterior distribution of the model parameters p and f as well as the type assigned to each lineage conditional to the observations.

Our Bayesian model can be viewed as the following two-step generative process:

1. Each lineage i is assigned to a progenitor type t_i drawn independently from a categorical distribution with frequencies f .
2. The occupancy vector S_{ij} of each lineage i at each layer j is drawn from a binomial distribution $\text{Binomial}(p_{t_i}^{(j)}; N_{\max})$ where $N_{\max}=20$ is the maximum number of cells that can occupy each layer.

The likelihood of a given lineage assignment and count matrix can be written as

$$P(t_{1,\dots,K}, S | p_{1,\dots,K}, f_{1:K}) = \left(\prod_{t=1}^K f_t^{n_t} \right) \prod_{i=1}^N \prod_{j \in \text{layers}} \left[p_{t_i}^{S_{ij}} \left(1 - p_{t_i}^{N_{\max} - S_{ij}} \right) \right]^{\sigma_{ij}},$$

where we introduced the binary variable σ_{ij} to denote whether the matrix element S_{ij} is included in the likelihood (in which case $\sigma_{ij} = 1$) or not ($\sigma_{ij} = 0$). In particular, the selection variable σ_{ij} for each lineage was set in such a way to exclude for each lineage the most superficial empty layers not followed by an occupied layer. The corresponding zero

counts in the matrix S might be spurious due to external processes stopping the lineage at early stages.

To perform Bayesian inference, we used Dirichlet priors on the relative frequencies and Beta distributions as priors on the occupancy probabilities p_1, \dots, K . To draw samples of model parameters and progenitor types from the posterior distribution we implemented a Gibbs sampler (Diana, 2019) which combines data likelihood and prior distributions to explore the parameter space efficiently. In order to draw statistical samples of the number of classes K we employed the Dirichlet process prior technique which allows us to remove existing classes or introduce new ones when assigning lineages to classes within the Gibbs sampler (Diana, 2019).

Sequential stochastic modelling of lineage generation

Probabilistic models 1 and 2 simulated 100 progenitors undergoing cell generation sequentially, following the in vivo inside-out pattern (*Figure 7—source data 2*). In each layer, in silico progenitors took a number of stochastic decisions for neuron generation; at each decision, a new neuron could be generated, or alternatively, the chance could be skipped without neuron generation. Sequential generation of neurons thus used the following parameters. Number of opportunities per layer was set randomly and could vary between a minimum of one and a maximum equal to the maximum number of cells found for that layer in any single experimental lineage across our three types of experiments. This parameter establishes the number of stochastic decisions available to the progenitor and reflects the size of the temporal “window” within which a progenitor can generate neurons for a given layer. Probability of cell generation, also layer-specific, gave the likelihood that a neuron is actually generated at each decision point. Simulations were repeated 100 times. Lineages smaller than three cells or larger than 12 cells were discarded from analysis.

For each model, the set of laminar division probabilities was adjusted to fit the experimental data regarding clonal size and laminar fractions of cells. Model 1 used a unique progenitor, i.e. a single set of laminar division probabilities. Model 2 incorporated an additional population and was fit by varying both the relative size of the two populations and the values of their division probabilities, including how probabilities varied across layers.

Quantification and statistical analysis

Cell distributions and clonal spatial configuration. In all the experiments, brain sections were sequentially analyzed in rostral-to-caudal order and PC clones throughout the entire neocortex were identified as sparse, spatially separated cell clusters. The boundaries between cortical layers were traced based on nuclear (DAPI) staining and the laminar position of each cell was recorded accordingly. PC clones were classified as translaminar, infragranular and supragranular clones according to the laminar position of the neurons belonging to each clone. Cortical areas were identified based on the reference atlas of adult mouse brain (Allen Brain Atlas; <http://www.brain-map.org>). In $Emx1^{CreER};MADM^{TG/GT}$ experiments, lineages derived from symmetric divisions (defined as lineages with three or more cells expressing each reporter) were excluded. In the $Emx1^{CreER};RCL-Gfp$ experiments, lineages derived from symmetric divisions (defined as lineages containing more than twelve neurons) were excluded. Lineages containing one or two cells were also excluded in $Emx1^{CreER};MADM^{TG/GT}$ and $Emx1^{CreER};RCL-Gfp$ experiments. In retroviral experiments, one-cell and two-cell clones were considered separately in E9.5, E10.5, E11.5 and E12.5 experiments. In E14.5 experiments, 2-cell lineages were considered, while one-cell clones were not quantified. This is due to the fact that, unlike at earlier time points,

RGCs may be undergoing their last neurogenic division at this stage, and thus lineages with a minimum of two cells may be derived from targeted apical progenitor cells.

Pyramidal cell types. Brain sections were stained for markers of cortical projection neuron identity and classified based on the relative expression of the transcription factors Ctip2 and Satb2 in four main subtypes: Cortico-cortical (CCPN), sub-cerebral (SCPN), cortico-thalamic (CThPN) and heterogeneous (HPN) projection neurons. This last type was defined as layer V cells expressing both Ctip2 and Satb2 markers, which have been recently described as a distinct identity³³. Images were captured using a confocal microscope and analyzed using a custom algorithm written in MATLAB (Mathworks). In brief, cell nuclei were segmented using the disk morphological function based on size and thresholds of fluorescence intensity over background. Cells were categorized as expressing high or low levels of the transcription factors Ctip2 and Satb2 and further subclassified as CCPN (Ctip2^{Low}/Satb2^{High}), SCPN (Ctip2^{High}/Satb2^{Low}) or HPN (Ctip2^{High}/Satb2^{High}), based on the combination of marker expression. To distinguish between CThPN from CCPN in layer VI we used the following criteria: CCPN (Ctip2^{Low}/Satb2^{High} OR Ctip2^{Low}/Satb2^{Low}) or CThPN (Ctip2^{High}/Satb2^{Low}). This allowed for the subclassification of layer V and layer VI cells based on the same set of markers. We verified these criteria by staining brain sections for the transcription factor Tle4 (*Figure 5–figure supplement 2*), a well-established specific marker of cortical CThPN identity (Molyneaux et al., 2015). Layer VI cells expressing high levels of both transcription factors were not classified, and lineages containing those cells were excluded from the quantification.

Laminar ratios. To quantify the actual densities of PCs in different cortical layers, *Neurod6*^{Cre} mice were crossed with Fucci2a reporter mice. The density of labeled red nuclei in each cortical layer was quantified from 5 representative serial sections of the

somatosensory and visual cortex. Z-stacks were then 3d reconstructed and quantified using Imaris 8.1.2 (Bitplane; RRID:SCR_007370).

Statistical tests. Error bars in all graphs indicate standard deviation (std) unless otherwise stated in the legends. Comparisons of distributions over fractions of a total (e.g., Figure 6e,f) were analyzed using Fisher's exact test or Chi-square test. Average clonal size between lineages analyzed at P2 and P21 were analyzed using Mann-Whitney U-test. All statistical tests are specified in the figure legends.

Acknowledgments

We thank I. Andrew and S.E. Bae for excellent technical assistance, F. Gage for plasmids, and K. Nave for *Nex-Cre* mice. We thank members of the Marín and Rico laboratories for stimulating discussions and ideas. This work was supported by grants from the European Research Council (ERC-2017-AdG 787355 to O.M. and ERC-2016-CoG 725780 to S.H.). F.K.W. was supported by an EMBO postdoctoral fellowship and is currently a Marie Skłodowska-Curie Fellow from the European Commission under the H2020 Programme, and R.B. received support from FWF Lise-Meitner program (M 2416).

Competing interest

The authors declare no competing interests.

773 **References**

- 774 Beattie, R., Postiglione, M.P., Burnett, L.E., Laukoter, S., Streicher, C., Pauler, F.M., Xiao,
775 G., Klezovitch, O., Vasioukhin, V., Ghashghaei, T.H., and Hippenmeyer, S. (2017).
776 Mosaic Analysis with Double Markers reveals distinct sequential functions of Lgl1 in
777 neural stem cells. *Neuron* 94, 517-533 e513.
- 778 Blanquie, O., Liebmann, L., Hubner, C.A., Luhmann, H.J., and Sinning, A. (2017).
779 NKCC1-mediated GABAergic signaling promotes postnatal cell death in neocortical
780 Cajal-Retzius cells. *Cereb. Cortex* 27, 1644-1659.
- 781 Blaschke, A.J., Staley, K., and Chun, J. (1996). Widespread programmed cell death in
782 proliferative and postmitotic regions of the fetal cerebral cortex. *Development* 122, 1165-
783 1174.
- 784 Brodmann, K., and Gary, L.J. (2006). Brodmann's localization in the cerebral cortex: the
785 principles of comparative localisation in the cerebral cortex based on cytoarchitectonics
786 (New York: Springer).
- 787 Cepko, C.L., Ryder, E., Austin, C., Golden, J., Fields-Berry, S., and Lin, J. (2000). Lineage
788 analysis with retroviral vectors. *Methods Enzymol.* 327, 118-145.
- 789 Ciceri, G., Dehorter, N., Sols, I., Huang, Z.J., Maravall, M., and Marín, O. (2013).
790 Lineage-specific laminar organization of cortical GABAergic interneurons. *Nat. Neurosci.*
791 16, 1199-1210.
- 792 Desai, A.R., and McConnell, S.K. (2000). Progressive restriction in fate potential by neural
793 progenitors during cerebral cortical development. *Development* 127, 2863-2872.
- 794 Diana, G. (2019). SampLin. <https://bitbucket.org/giovannidiana/samplin/src/master>
795 e46d5a1.

796 Eckler, M.J., Nguyen, T.D., McKenna, W.L., Fastow, B.L., Guo, C., Rubenstein, J.L.R.,
 797 and Chen, B. (2015). Cux2-positive radial glial cells generate diverse subtypes of
 798 neocortical projection neurons and macroglia. *Neuron* 86, 1100-1108.

799 Franco, S.J., Gil-Sanz, C., Martinez-Garay, I., Espinosa, A., Harkins-Perry, S.R., Ramos,
 800 C., and Muller, U. (2012). Fate-restricted neural progenitors in the mammalian cerebral
 801 cortex. *Science* 337, 746-749.

802 Franco, S.J., and Muller, U. (2013). Shaping our minds: stem and progenitor cell diversity
 803 in the mammalian neocortex. *Neuron* 77, 19-34.

804 Gal, J.S., Morozov, Y.M., Ayoub, A.E., Chatterjee, M., Rakic, P., and Haydar, T.F. (2006).
 805 Molecular and morphological heterogeneity of neural precursors in the mouse neocortical
 806 proliferative zones. *J. Neurosci.* 26, 1045-1056.

807 Gao, P., Postiglione, M.P., Krieger, T.G., Hernandez, L., Wang, C., Han, Z., Streicher, C.,
 808 Papusheva, E., Insolera, R., Chugh, K., *et al.* (2014). Deterministic progenitor behavior
 809 and unitary production of neurons in the neocortex. *Cell* 159, 775-788.

810 García-Moreno, F., and Molnar, Z. (2015). Subset of early radial glial progenitors that
 811 contribute to the development of callosal neurons is absent from avian brain. *Proc. Natl.*
 812 *Acad. Sci. USA* 112, E5058-5067.

813 Gaspard, N., Bouchet, T., Hourez, R., Dimidschstein, J., Naeije, G., van den Aamele, J.,
 814 Espuny-Camacho, I., Herpoel, A., Passante, L., Schiffmann, S.N., *et al.* (2008). An
 815 intrinsic mechanism of corticogenesis from embryonic stem cells. *Nature* 455, 351-357.

816 Gil-Sanz, C., Espinosa, A., Fregoso, S.P., Bluske, K.K., Cunningham, C.L., Martinez-
 817 Garay, I., Zeng, H., Franco, S.J., and Muller, U. (2015). Lineage tracing using Cux2-Cre
 818 and Cux2-CreERT2 mice. *Neuron* 86, 1091-1099.

819 Goebbels, S., Bormuth, I., Bode, U., Hermanson, O., Schwab, M.H., and Nave, K.A.
 820 (2006). Genetic targeting of principal neurons in neocortex and hippocampus of NEX-Cre
 821 mice. *Genesis* 44, 611-621.

822 Gomes, F.L., Zhang, G., Carbonell, F., Correa, J.A., Harris, W.A., Simons, B.D., and
 823 Cayouette, M. (2011). Reconstruction of rat retinal progenitor cell lineages in vitro reveals
 824 a surprising degree of stochasticity in cell fate decisions. *Development* 138, 227-235.

825 Gotz, M., and Huttner, W.B. (2005). The cell biology of neurogenesis. *Nat. Rev. Mol. Cell*
 826 *Biol.* 6, 777-788.

827 Greig, L.C., Woodworth, M.B., Galazo, M.J., Padmanabhan, H., and Macklis, J.D. (2013).
 828 Molecular logic of neocortical projection neuron specification, development and diversity.
 829 *Nat. Rev. Neurosci.*

830 Guo, C., Eckler, M.J., McKenna, W.L., McKinsey, G.L., Rubenstein, J.L., and Chen, B.
 831 (2013). Fezf2 expression identifies a multipotent progenitor for neocortical projection
 832 neurons, astrocytes, and oligodendrocytes. *Neuron* 80, 1167-1174.

833 Harb, K., Magrinelli, E., Nicolas, C.S., Lukianets, N., Frangeul, L., Pietri, M., Sun, T.,
 834 Sandoz, G., Grammont, F., Jabaudon, D., *et al.* (2016). Area-specific development of
 835 distinct projection neuron subclasses is regulated by postnatal epigenetic modifications.
 836 *Elife* 5, e09531.

837 Haubensak, W., Attardo, A., Denk, W., and Huttner, W.B. (2004). Neurons arise in the
 838 basal neuroepithelium of the early mammalian telencephalon: a major site of neurogenesis.
 839 *Proc. Natl. Acad. Sci. USA* 101, 3196-3201.

840 Haydar, T.F., Ang, E., Jr., and Rakic, P. (2003). Mitotic spindle rotation and mode of cell
 841 division in the developing telencephalon. *Proc. Natl. Acad. Sci. USA* 100, 2890-2895.

842 He, J., Zhang, G., Almeida, A.D., Cayouette, M., Simons, B.D., and Harris, W.A. (2012).
 843 How variable clones build an invariant retina. *Neuron* 75, 786-798.
 844 Hippenmeyer, S., Youn, Y.H., Moon, H.M., Miyamichi, K., Zong, H., Wynshaw-Boris, A.,
 845 and Luo, L. (2010). Genetic mosaic dissection of *Lis1* and *Ndel1* in neuronal migration.
 846 *Neuron* 68, 695-709.
 847 Jabaudon, D. (2017). Fate and freedom in developing neocortical circuits. *Nat. Commun.*
 848 8, 16042.
 849 Kaplan, E.S., Ramos-Laguna, K.A., Mihalas, A.B., Daza, R.A.M., and Hevner, R.F.
 850 (2017). Neocortical Sox9+ radial glia generate glutamatergic neurons for all layers, but
 851 lack discernible evidence of early laminar fate restriction. *Neural Development* 12, 14.
 852 Kessaris, N., Fogarty, M., Iannarelli, P., Grist, M., Wegner, M., and Richardson, W.D.
 853 (2006). Competing waves of oligodendrocytes in the forebrain and postnatal elimination of
 854 an embryonic lineage. *Nat. Neurosci.* 9, 173-179.
 855 Kriegstein, A.R., and Gotz, M. (2003). Radial glia diversity: a matter of cell fate. *Glia* 43,
 856 37-43.
 857 Lange, C., Huttner, W.B., and Calegari, F. (2009). Cdk4/cyclinD1 overexpression in neural
 858 stem cells shortens G1, delays neurogenesis, and promotes the generation and expansion of
 859 basal progenitors. *Cell Stem Cell* 5, 320-331.
 860 Leone, D.P., Srinivasan, K., Chen, B., Alcamo, E., and McConnell, S.K. (2008). The
 861 determination of projection neuron identity in the developing cerebral cortex. *Curr. Opin.*
 862 *Neurobiol.* 18, 28-35.
 863 Lodato, S., and Arlotta, P. (2015). Generating neuronal diversity in the mammalian
 864 cerebral cortex. *Annu. Rev. Cell. Dev. Biol.* 31, 699-720.

865 Lui, J.H., Hansen, D.V., and Kriegstein, A.R. (2011). Development and evolution of the
 866 human neocortex. *Cell* *146*, 18-36.

867 Luskin, M.B., Pearlman, A.L., and Sanes, J.R. (1988). Cell lineage in the cerebral cortex of
 868 the mouse studied in vivo and in vitro with a recombinant retrovirus. *Neuron* *1*, 635-647.

869 Malatesta, P., Hartfuss, E., and Gotz, M. (2000). Isolation of radial glial cells by
 870 fluorescent-activated cell sorting reveals a neuronal lineage. *Development* *127*, 5253-5263.

871 Mihalas, A.B., Elsen, G.E., Bedogni, F., Daza, R.A.M., Ramos-Laguna, K.A., Arnold, S.J.,
 872 and Hevner, R.F. (2016). Intermediate progenitor cohorts differentially generate cortical
 873 layers and require Tbr2 for timely acquisition of neuronal subtype identity. *Cell Reports*
 874 *16*, 92-105.

875 Mihalas, A.B., and Hevner, R.F. (2018). Clonal analysis reveals laminar fate multipotency
 876 and daughter cell apoptosis of mouse cortical intermediate progenitors. *Development* *145*.

877 Miyata, T., Kawaguchi, A., Okano, H., and Ogawa, M. (2001). Asymmetric inheritance of
 878 radial glial fibers by cortical neurons. *Neuron* *31*, 727-741.

879 Miyata, T., Kawaguchi, A., Saito, K., Kawano, M., Muto, T., and Ogawa, M. (2004).
 880 Asymmetric production of surface-dividing and non-surface-dividing cortical progenitor
 881 cells. *Development* *131*, 3133-3145.

882 Mizutani, K.I., Yoon, K., Dang, L., Tokunaga, A., and Gaiano, N. (2007). Differential
 883 Notch signalling distinguishes neural stem cells from intermediate progenitors. *Nature* *449*,
 884 351-355.

885 Molyneaux, B.J., Arlotta, P., Menezes, J.R., and Macklis, J.D. (2007). Neuronal subtype
 886 specification in the cerebral cortex. *Nat. Rev. Neurosci.* *8*, 427-437.

887 Molyneaux, B.J., Goff, L.A., Brettler, A.C., Chen, H.H., Hrvatin, S., Rinn, J.L., and
 888 Arlotta, P. (2015). DeCoN: genome-wide analysis of in vivo transcriptional dynamics
 889 during pyramidal neuron fate selection in neocortex. *Neuron* 85, 275-288.
 890 Mort, R.L., Ford, M.J., Sakaue-Sawano, A., Lindstrom, N.O., Casadio, A., Douglas, A.T.,
 891 Keighren, M.A., Hohenstein, P., Miyawaki, A., and Jackson, I.J. (2014). Fucci2a: a
 892 bicistronic cell cycle reporter that allows Cre mediated tissue specific expression in mice.
 893 *Cell Cycle* 13, 2681-2696.
 894 Noctor, S.C., Flint, A.C., Weissman, T.A., Dammerman, R.S., and Kriegstein, A.R.
 895 (2001). Neurons derived from radial glial cells establish radial units in neocortex. *Nature*
 896 409, 714-720.
 897 Noctor, S.C., Martinez-Cerdeno, V., Ivic, L., and Kriegstein, A.R. (2004). Cortical neurons
 898 arise in symmetric and asymmetric division zones and migrate through specific phases.
 899 *Nat. Neurosci.* 7, 136-144.
 900 Oberst, P., Fievre, S., Baumann, N., Concetti, C., Bartolini, G., and Jabaudon, D. (2019).
 901 Temporal plasticity of apical progenitors in the developing mouse neocortex. *Nature* 573,
 902 370-374.
 903 Pilaz, L.J., Patti, D., Marcy, G., Ollier, E., Pfister, S., Douglas, R.J., Betizeau, M., Gautier,
 904 E., Cortay, V., Doerflinger, N., *et al.* (2009). Forced G1-phase reduction alters mode of
 905 division, neuron number, and laminar phenotype in the cerebral cortex. *Proc. Natl. Acad.*
 906 *Sci. USA* 106, 21924-21929.
 907 Pimeisl, I.M., Tanriver, Y., Daza, R.A., Vauti, F., Hevner, R.F., Arnold, H.H., and Arnold,
 908 S.J. (2013). Generation and characterization of a tamoxifen-inducible Eomes(CreER)
 909 mouse line. *Genesis* 51, 725-733.

910 Pla, R., Borrell, V., Flames, N., and Marín, O. (2006). Layer acquisition by cortical
911 GABAergic interneurons is independent of Reelin signaling. *J. Neurosci.* 26, 6924-6934.

912 Pollen, A.A., Nowakowski, T.J., Shuga, J., Wang, X., Leyrat, A.A., Lui, J.H., Li, N.,
913 Szpankowski, L., Fowler, B., Chen, P., *et al.* (2014). Low-coverage single-cell mRNA
914 sequencing reveals cellular heterogeneity and activated signaling pathways in developing
915 cerebral cortex. *Nat. Biotechnol.* 32, 1053-1058.

916 Price, J., and Thurlow, L. (1988). Cell lineage in the rat cerebral cortex: a study using
917 retroviral-mediated gene transfer. *Development* 104, 473-482.

918 Rakic, P. (1974). Neurons in rhesus monkey visual cortex: systematic relation between
919 time of origin and eventual disposition. *Science* 183, 425-427.

920 Rakic, P. (1988). Specification of cerebral cortical areas. *Science* 241, 170-176.

921 Rakic, P., Bourgeois, J.P., and Goldman-Rakic, P.S. (1994). Synaptic development of the
922 cerebral cortex: implications for learning, memory, and mental illness. *Prog. Brain Res.*
923 102, 227-243.

924 Reid, C.B., Liang, I., and Walsh, C. (1995). Systematic widespread clonal organization in
925 cerebral cortex. *Neuron* 15, 299-310.

926 Roccio, M., Schmitter, D., Knobloch, M., Okawa, Y., Sage, D., and Lutolf, M.P. (2013).
927 Predicting stem cell fate changes by differential cell cycle progression patterns.
928 *Development* 140, 459-470.

929 Shen, Q., Wang, Y., Dimos, J.T., Fasano, C.A., Phoenix, T.N., Lemischka, I.R., Ivanova,
930 N.B., Stifani, S., Morrissey, E.E., and Temple, S. (2006). The timing of cortical
931 neurogenesis is encoded within lineages of individual progenitor cells. *Nat. Neurosci.* 9,
932 743-751.

933 Sousa, V.H., Miyoshi, G., Hjerling-Leffler, J., Karayannis, T., and Fishell, G. (2009).
934 Characterization of Nkx6-2-derived neocortical interneuron lineages. *Cereb. Cortex* 19
935 *Suppl 1*, i1-10.

936 Tan, S.S., Kalloniatis, M., Sturm, K., Tam, P.P., Reese, B.E., and Faulkner-Jones, B.
937 (1998). Separate progenitors for radial and tangential cell dispersion during development
938 of the cerebral neocortex. *Neuron* 21, 295-304.

939 Tashiro, A., Sandler, V.M., Toni, N., Zhao, C., and Gage, F.H. (2006). NMDA-receptor-
940 mediated, cell-specific integration of new neurons in adult dentate gyrus. *Nature* 442, 929-
941 933.

942 Tasic, B., Yao, Z., Graybuck, L.T., Smith, K.A., Nguyen, T.N., Bertagnolli, D., Goldy, J.,
943 Garren, E., Economo, M.N., Viswanathan, S., *et al.* (2018). Shared and distinct
944 transcriptomic cell types across neocortical areas. *Nature* 563, 72-78.

945 Taverna, E., Gotz, M., and Huttner, W.B. (2014). The cell biology of neurogenesis: toward
946 an understanding of the development and evolution of the neocortex. *Annu. Rev. Cell.*
947 *Dev. Biol.* 30, 465-502.

948 Telley, L., Govindan, S., Prados, J., Stevant, I., Nef, S., Dermitzakis, E., Dayer, A., and
949 Jabaudon, D. (2016). Sequential transcriptional waves direct the differentiation of newborn
950 neurons in the mouse neocortex. *Science* 351, 1443-1446.

951 Vitali, I., Fievre, S., Telley, L., Oberst, P., Bariselli, S., Frangeul, L., Baumann, N.,
952 McMahon, J.J., Klingler, E., Bocchi, R., *et al.* (2018). Progenitor hyperpolarization
953 regulates the sequential generation of neuronal subtypes in the developing neocortex. *Cell*
954 174, 1264-1276.

955 Walsh, C., and Cepko, C.L. (1988). Clonally related cortical cells show several migration
956 patterns. *Science* 241, 1342-1345.

957 Walsh, C., and Cepko, C.L. (1992). Widespread dispersion of neuronal clones across
 958 functional regions of the cerebral cortex. *Science* 255, 434-440.
 959 Wang, X., Tsai, J.W., Imai, J.H., Lian, W.N., Vallee, R.B., and Shi, S.H. (2009).
 960 Asymmetric centrosome inheritance maintains neural progenitors in the neocortex. *Nature*
 961 461, 947-955.
 962 Wong, F.K., Bercsenyi, K., Sreenivasan, V., Portales, A., Fernandez-Otero, M., and Marin,
 963 O. (2018). Pyramidal cell regulation of interneuron survival sculpts cortical networks.
 964 *Nature* 557, 668-673.
 965 Zong, H., Espinosa, J.S., Su, H.H., Muzumdar, M.D., and Luo, L. (2005). Mosaic analysis
 966 with double markers in mice. *Cell* 121, 479-492.
 967

Figure legends

Figure 1. Identification of pyramidal cell lineages with low-titer conditional reporter retroviruses. **(a)** Experimental paradigm. **(b)** Schematic representation of the expected labeling outcomes in retroviral lineage tracing experiments. **(c–c’)** Serial 100µm coronal sections through the telencephalon of a P21 *Neurod6*^{Cre/+} mouse infected with low-titer conditional reporter retroviruses at E11.5. Lineages (L) 1 and 3 are shown at high magnification in c’ and c’’, respectively. Dashed lines define external brain boundaries and cortical layers. The schemas collapse lineages spanning across several sections into a single diagram. **(d)** Quantification of the number of PCs per lineage in P21 *Neurod6*^{Cre/+} mice infected with conditional reporter retroviruses at different embryonic stages. Lineages smaller than three cells were excluded. Boxes show median and inter-quartile distance, whiskers correspond to minimum and maximum values. Colored dots show individual clonal size values. **(e)** Quantification of the fraction of cortical lineages containing one, two or three or more neurons in P21 *Neurod6*^{Cre/+} mice infected with conditional reporter retroviruses at different embryonic stages. n = 13 lineages in 3 animals at E9.5; 21 lineages in 3 animals at E10.5; 64 lineages in 5 animals at E11.5; 166 lineages in 7 animals at E12.5; 32 lineages in 4 animals at E14.5. I–VI, cortical layers I to VI; H, hippocampus area; M1, primary motor cortex; RSD, retrosplenial cortex; S1, primary somatosensory cortex. Scale bars equal 100 µm (c) and 300 µm (c’ and c’’). Data used for quantitative analyses as well as the numerical data that are represented in graphs are available in [Figure 1—source data 1](#) and [Figure 1—source data 2](#). See also [Figure 1—figure supplement 1](#).

Figure 1—source data 1. E9.5-E14.5 *Neurod6*^{Cre/+} retroviral lineages analyzed at P21.

Figure 1—source data 2. Summary of numerical data that are represented in graphs.

Figure 1—figure supplement 1. Sparse labeling of neuronal clones with low-titer

retroviral infection. **(a)** Experimental paradigm. **(b)** Fraction of neuron clusters containing

cells labeled with one or two reporters. **(c,d)** Serial 100µm coronal sections through the

telencephalon of P21 *Neurod6*^{Cre/+} mice infected with low-titer conditional reporter

retroviruses at E9.5 (c) and E10.5 (d). Example images illustrating each lineage correspond

to sequential sections of the same brain. The high magnification pictures shown in (c' and

c'') and (d' and d'') correspond to the clones shown in (c) and (d), respectively. n = 296

lineages in 19 animals across all ages. Scale bars equal 100 µm (c and d) and 300 µm (c',

c'', d' and d''). Data used for quantitative analyses as well as the numerical data that are

represented in graphs are available in [Figure 1—source data 1](#) and [Figure 1—source data](#)

[2](#).

Figure 2. Retroviral-based lineage tracing reveals diverse lineage outcomes. **(a)**

Experimental paradigm. The bottom panel illustrates the expected labeling outcome

following retroviral infection of an RGC undergoing a neurogenic cell division in which

the viral integration occurs in the self-renewing RGC. **(b–d)** Serial 100µm coronal sections

through the cortex of P21 *Neurod6*^{Cre/+} mice infected with low-titer conditional reporter

retroviruses at E12.5. The images show examples of translaminar (b), deep-layer restricted

(c) and superficial-layer restricted (d) lineages containing three or more cells. Dashed lines

define external brain boundaries and cortical layers. The schemas collapse lineages

spanning across several sections into a single diagram. Example images illustrating each

lineage correspond to sequential sections of the same brain. **(e)** Quantification of the

fraction of translaminar, deep- and superficial-layer restricted lineages containing three or more cells, and clonal size. Boxes show median and inter-quartile distance, whiskers correspond to minimum and maximum values. Colored dots show individual clonal size values. **(f)** Expected labeling outcome following retroviral infection of an RGC undergoing a neurogenic cell division in which the viral integration occurs in an IPC (indirect neurogenesis). **(g,h)** Coronal sections through the cortex of P21 *Neurod6*^{Cre/+} mice infected with low-titer conditional reporter retroviruses at E12.5. The images show examples of superficial and deep layer-restricted two-cell clones. **(i)** Quantification of the fraction of translaminar, deep and superficial layer-restricted two-cell lineages. **(j)** Expected labeling outcome following retroviral infection of an RGC undergoing a neurogenic cell division in which the viral integration occurs in a postmitotic neuron (direct neurogenesis). **(k,l)** Coronal sections through the cortex of P21 *Neurod6*^{Cre/+} mice infected with low-titer conditional reporter retroviruses at E12.5. The images show examples of superficial and deep layer-restricted single-cell clones. **(m)** Laminar distribution of single-cell clones. n = 166 lineages in 7 animals. Scale bar equals 100 μm. Data used for quantitative analyses as well as the numerical data that are represented in graphs are available in [Figure 2—source data 1](#) and [Figure 1—source data 2](#). See also [Figure 2—figure supplement 1](#).

Figure 2—source data 1. E12.5 *Neurod6*^{Cre/+} retroviral lineages analyzed at P21.

Figure 2—figure supplement 1. **(a)** Experimental paradigm. The bottom panel illustrates the expected labeling outcome of the division of a Tbr2+ intermediate progenitor cell. **(b,c)** 100μm coronal sections through the cortex of P21 *Tbr2*^{CreERT2/+}; *RCL-Gfp* mice treated with low-dose tamoxifen at E12.5. The images show examples of deep and superficial layer-restricted Tbr2-derived lineages. The schemas collapse lineages spanning across several

sections into a single diagram. **(d)** Laminar fraction of translaminar, deep and superficial layer-restricted Tbr2-derived lineages. n = 72 lineages in 12 animals. Scale bar equals 100 μ m. Data used for quantitative analyses as well as the numerical data that are represented in graphs are available in [Figure 2—figure supplement 1—source data 1](#) and [Figure 1—source data 2](#).

Figure 2— figure supplement 1—source data 1. E12.5 *Tbr2*^{CreERT2/+}; *RCL-Gfp* lineages analyzed at P21.

Figure 3. Lineage tracing using MADM identifies a small fraction of deep layer-restricted cortical lineages. **(a)** Experimental paradigm. The bottom panel illustrates the expected labeling outcome of a neurogenic RGC division following inducible MADM-based lineage tracing in which two subclones are labeled with different reporters. **(b,c)** Serial 100 μ m coronal sections through the cortex of P21 *Emx1-CreER*^{T2}; MADM^{TG/GT} mice treated with tamoxifen at E12.5. The images show examples of translaminar **(b)** and deep layer-restricted **(c)** lineages. The schemas collapse lineages spanning across several sections into a single diagram. Example images illustrating each lineage correspond to sequential sections of the same brain. **(d)** Quantification of the fraction of translaminar, deep and superficial layer-restricted lineages, and clonal size in MADM lineages derived from a neurogenic (asymmetric) RGC division. Boxes show median and inter-quartile distance, whiskers correspond to minimum and maximum values. Colored dots show individual clonal size values. n = 106 neurogenic lineages in 28 animals. Scale bar equals 100 μ m. Data used for quantitative analyses as well as the numerical data that are represented in

graphs are available in [Figure 3—source data 1](#) and [Figure 1—source data 2](#). See also [Figure 3—figure supplement 1](#).

Figure 3—source data 1. E12.5 *Emx1-CreER^{T2}*;MADM^{TG/GT} lineages analyzed at P21.

Figure 3—figure supplement 1. Large MADM subclones reveal a small fraction of

artifactual superficial layer-restricted lineages in the retroviral experiments. (a)

Experimental paradigm. The bottom panel illustrates the expected labeling outcome of a

neurogenic RGC division following inducible MADM-based lineage tracing in which two

subclones are labeled with different reporters. (b) Serial 100µm coronal sections through

the cortex of a P21 *Emx1-CreER^{T2}*;MADM^{TG/GT} mouse treated with tamoxifen at E12.5.

The images show an example of a translaminal lineage containing a large superficial layer-

restricted sub-clone. The schemas collapse lineages spanning across several sections into a

single diagram. Example images correspond to sequential sections of the same brain. (c)

Relative fraction of translaminal, deep and superficial layer restricted MADM large sub-

clones. n = 196 sub-lineages in 28 animals. Scale bar equals 100 µm. Data used for

quantitative analyses as well as the numerical data that are represented in graphs are

available in [Figure 3—source data 1](#) and [Figure 1—source data 2](#).

Figure 4. A fraction of early-quiescent cortical progenitors generates superficial layer-

restricted lineages. (a) Experimental paradigm. The bottom panel illustrates the expected

labeling outcome of a neurogenic RGC division following inducible conditional reporter

lineage tracing in *Emx1-CreER^{T2}*;RCL-Gfp mice. (b–d) Serial 100µm coronal sections

through the cortex of P21 *Emx1-CreER^{T2}*;RCL-Gfp mice treated with low-dose tamoxifen

at E12.5. The images show examples of translaminar (b), deep layer-restricted (c) and superficial layer-restricted (d) lineages. The schemas collapse lineages spanning across several sections in a single diagram. Example images illustrating each lineage correspond to sequential sections of the same brain. (e) Quantification of the fraction of translaminar, deep and superficial layer-restricted lineages, and clonal size in inducible conditional reporter lineage tracing experiments. Boxes show median and inter-quartile distance, whiskers correspond to minimum and maximum values. Colored dots show individual clonal size values. n = 260 neurogenic lineages in 25 animals. Scale bar equals 100 μ m. Data used for quantitative analyses as well as the numerical data that are represented in graphs are available in [Figure 4—source data 1](#) and [Figure 1—source data 2](#). See also [Figure 4—figure supplement 1](#).

Figure 4—source data 1. E12.5 *Emx1-CreER^{T2};RCL-Gfp* lineages analyzed at P21.

Figure 4—figure supplement 1. Superficial layer-restricted lineages in the murine cerebral cortex. (a–d) Serial 100 μ m coronal sections through the cortex of P21 *Emx1-CreER^{T2};RCL-Gfp* mice treated with tamoxifen at E12.5. The images show examples of superficial lineages. The schemas collapse lineages spanning across several sections into a single diagram. Example images illustrating each lineage correspond to sequential sections of the same brain. Scale bar equals 150 μ m.

Figure 5. Translaminar lineages adopt very heterogeneous configurations. (a) Experimental paradigm. The bottom panel illustrates the expected labeling outcome of a neurogenic RGC division following inducible conditional reporter lineage tracing in *Emx1-*

1103 *CreER^{T2};RCL-Gfp* mice. **(b–d)** Serial 100 μ m coronal sections through the cortex of P21
 1104 *Emx1-CreER^{T2};RCL-Gfp* mice treated with low-dose tamoxifen at E12.5. The images show
 1105 examples of translaminar lineages with various laminar configurations observed in the
 1106 somatosensory (b and c) and visual cortex (d). The schemas collapse lineages spanning
 1107 across several sections into a single diagram. Example images illustrating each lineage
 1108 correspond to sequential sections of the same brain. **(e)** Clonal size distribution of
 1109 translaminar, deep and superficial layer-restricted lineages. **(f)** Relative frequency
 1110 (expressed as percentage over the total number of lineages) of the different laminar
 1111 configurations (green and grey schemas) in inducible conditional reporter lineage tracing
 1112 experiments. **(g)** Relative abundance of PCs in superficial and deep layers from
 1113 translaminar lineages containing cells in every layer. Lineages are represented as circles in
 1114 a bi-dimensional space, indicating the number of cells in superficial versus deep layers.
 1115 The size of the circle indicates the number of lineages that shown a particular
 1116 configuration. Green shapes schematically represent lineage configurations. A rectangular
 1117 shape illustrates lineages with a balanced number of superficial and deep PCs; triangular
 1118 shapes represent configurations of lineages biased towards superficial or deep layer
 1119 neurons. **(h)** Fraction of translaminar lineages with neurons in every layer containing one,
 1120 two, three or four subclasses of PCs. n = 260 neurogenic lineages in 25 animals. CCPN,
 1121 cortico-cortical projection neuron; SCPN, subcortical projection neuron; HPN,
 1122 heterogeneous projection neuron; CThPN, cortico-thalamic projection neuron. Scale bar
 1123 equals 100 μ m. Data used for quantitative analyses as well as the numerical data that are
 1124 represented in graphs are available in [Figure 4—source data 1](#), [Figure 5—source data 1](#)
 1125 and [Figure 1—source data 2](#). See also [Figure 5—figure supplements 1-3](#).

Figure 5—source data 1. Projection neuron markers in E12.5 *Emx1-CreER^{T2};RCL-Gfp* all-layer lineages analyzed at P21.

Figure 5—figure supplement 1. Retrovirus and MADM labeled lineages exhibit a diversity of laminar configurations. **(a)** Experimental paradigm. The right panel illustrates the expected labeling outcome following retroviral infection of an RGC undergoing a neurogenic cell division in which the viral integration occurs in the self-renewing RGC. **(b,c)** Serial 100µm coronal sections through the cortex of P21 *Neurod6^{Cre/+}* mice infected with low-titer conditional reporter retroviruses at E12.5. The images show examples of layer V (b) and layer 6 (c) skipping lineages. Dashed lines define cortical layers. The schemas collapse lineages spanning across several sections into a single diagram. Example images illustrating each lineage correspond to sequential sections of the same brain. **(d)** Clonal size distribution of translaminar, deep and superficial layer-restricted lineages in retroviral experiments. **(e)** Relative frequency (expressed as percentage over the total number of lineages) of the different laminar configurations (green and grey schemas) in retroviral labeling experiments. **(f)** Experimental paradigm. The bottom panel illustrates the expected labeling outcome of a neurogenic RGC division following inducible MADM-based lineage tracing in which two subclones are labeled with different reporters. **(g,h)** Serial 100µm coronal sections through the cortex of P21 *Emx1-CreER^{T2};MADM^{TG/GT}* mice treated with tamoxifen at E12.5. The images show examples of layer IV and V restricted (g) and layer V skipping (h) lineages. The schemas collapse lineages spanning across several sections into a single diagram. Example images illustrating each lineage correspond to sequential sections of the same brain. **(i)** Clonal size distribution of translaminar, deep and superficial layer-restricted lineages in MADM experiments. **(j)** Relative frequency (expressed as percentage over the total number of lineages) of the different laminar

configurations (green and grey schemas) in MADM experiments. n = 73 lineages in 7 animals in retrovirus experiments (d and e); 106 neurogenic lineages in 28 animals in MADM experiments (I and j). Scale bar equals 100 μ m. Data used for quantitative analyses as well as the numerical data that are represented in graphs are available in [Figure 2—source data 1](#), [Figure 3—source data 1](#) and [Figure 1—source data 2](#).

Figure 5—figure supplement 2. Identification of pyramidal cell subclasses. **(a)** Experimental paradigm. The bottom panel illustrates the expected labeling outcome of a neurogenic RGC division following inducible conditional reporter lineage tracing in *Emx1-CreER^{T2};RCL-Gfp* mice. **(b)** Schematic of PC subclasses based on their laminar distribution and expression of specific markers. **c**, Fraction of Tle4+ and Tle4- layer VI PCs that express the transcription factors *Satb2* and *Ctip2*. Data are represented as mean \pm standard deviation. **(d)** Pyramidal cell subclasses in lineages traced in *Emx1-CreER^{T2};RCL-Gfp* mice. Neurons are classified according to their laminar allocation and expression of *Satb2* and *Ctip2*. n = 1,123 cells in two brains. CCPN, cortico-cortical projection neuron; SCPN, subcortical projection neuron; HPN, heterogeneous projection neuron; CThPN, cortico-thalamic projection neuron. Scale bar equals 35 μ m. Data used for quantitative analyses as well as the numerical data that are represented in graphs are available in [Figure 5—figure supplement 2—source data 1](#) and [Figure 1—source data 2](#).

Figure 5—figure supplement 2—source data 1. Projection neuron marker colocalization.

Figure 5—figure supplement 3. Subtle impact of pyramidal neuron cell death in final configurations of cortical neuron lineages. **(a)** Experimental paradigm. The bottom panel illustrates the expected labeling outcome of a neurogenic RGC division following

inducible conditional reporter lineage tracing in *Emx1-CreER^{T2};RCL-Gfp* mice. **(b)** Clonal size of cortical lineages at P2 and P21 following low-dose tamoxifen in *Emx1-CreER^{T2};RCL-Gfp* mice at E12.5. Data are represented as mean \pm standard deviation. **(c)** Clonal size distribution of cortical lineages at P2 and P21 in inducible conditional reporter lineage tracing experiments. **(d)** Quantification of the fraction of translaminal, deep and superficial layer-restricted lineages at P2 and P21 in inducible conditional reporter lineage tracing experiments. **(e)** Relative frequency (expressed as percentage over the total number of lineages) of the different laminar configurations (green and grey schemas) in inducible conditional reporter lineage tracing experiments at P2 and P21. At this age, layers II/II and IV cannot be reliably distinguished and are consequently considered as a single cortical layer. n = 110 neurogenic lineages in 11 animals. Data used for quantitative analyses as well as the numerical data that are represented in graphs are available in [Figure 4—source data 1](#), [Figure 5—figure supplement 3—source data 1](#) and [Figure 1—source data 2](#).

Figure 5—figure supplement 3—source data 1. E12.5 *Emx1-CreER^{T2};RCL-Gfp* lineages analyzed at P2.

Figure 6. A stochastic model of cortical neurogenesis. **(a)** Number of progenitor identities required to reproduce experimental lineage configurations inferred by Bayesian modeling. The y axis represents the fraction of simulations from a total of 4000 that demand a particular minimum number of progenitors. **(b)** Fraction of cells in each cortical layer (expressed as percentage of total) in experimental and modeled lineages. **(c)** Clonal size distribution in experimental and modeled lineages. **(d)** Spearman correlation (r) values for

the fraction of superficial and deep layer neurons in modeled lineages. Each dot represents an r value for one simulation. The green line shows the experimental value; the shadow area around the experimental data represents a 95% confidence interval for the experimental value. **(e)** Fraction of translaminal, deep and superficial layer-restricted lineages found experimentally and predicted by the model (expressed as percentage of all modeled lineages within a single simulation). Grey boxes represent variability among 4000 simulations; colored stars and lines show experimental values and 95% confidence intervals for experimental values ($p = 0.036$, Chi-square test). **(f)** Relative frequency (expressed as percentage over all modeled translaminal lineages within a single simulation) of laminar configurations in experimental and modeled translaminal lineages. Grey boxes represent variability among 4000 simulations; colored stars and lines show experimental values and 95% confidence intervals for experimental values ($p = 0.11$, Chi-square test). Histograms represent mean \pm standard deviation. Z-scores represent the distance between experimental and simulated results for each parameter, which is calculated as the difference between the averages of model and experimental data divided by the standard deviation within model simulations (see methods for details). $n = 103$ neurogenic lineages in the somatosensory of 25 animals. Data used for quantitative analyses as well as the numerical data that are represented in graphs are available in [Figure 4—source data 1](#) and [Figure 1—source data 2](#). See also [Figure 6—figure supplement 1](#).

Figure 6—figure supplement 1. Laminar densities of pyramidal neurons do not predict lineage structure. **(a)** Fraction of cells in each cortical layer (expressed as percentage of total) in experimental and bootstrapped lineages. **(b)** Clonal size distribution in experimental and bootstrapped lineages. **(c)** Spearman correlation (r) values for the fraction

of superficial and deep layer neurons in modeled lineages. Each dot represents an r value for one simulation. The green line shows the experimental value; the shadow area around the experimental data represents a 95% confidence interval for the experimental value. Histograms represent mean \pm standard deviation. Z-scores represent the distance between experimental and simulated results for each parameter, which is calculated as the difference between the averages of model and experimental data divided by the standard deviation within model simulations (see methods for details). $n = 103$ neurogenic lineages in the somatosensory cortex of 25 animals. Data used for quantitative analyses as well as the numerical data that are represented in graphs are available in [Figure 4—source data 1](#) and [Figure 1—source data 2](#).

Figure 7. A small number of progenitor identities underlie lineage diversity. **(a)** Schematic representation of a mathematical model of cortical neurogenesis in which two different progenitor identities are modeled (Model 2). Squares represent the maximum number of stochastic decisions performed by each progenitor for each cortical layer during *in silico* simulations. The odds of generating neurons for each chance are given by a probability value (P), which is unique for each layer and progenitor identity. The model runs 100 simulations with 100 progenitors. **(b)** Fraction of cells in each cortical layer (expressed as percentage of total) in experimental and modeled lineages. **(c)** Clonal size distribution in experimental and modeled lineages. **(d)** Spearman correlation (r) values for the fraction of superficial and deep layer neurons in modeled lineages. Each dot represents an r value for one simulation. The green line shows the experimental value; the shadow area around the experimental data represents a 95% confidence interval for the experimental value. **(e)**

Fraction of translaminal, deep and superficial layer-restricted lineages found experimentally and predicted by the model (expressed as percentage of all modeled lineages within a single simulation). Grey boxes represent variability among 100 simulations; colored stars and lines show experimental values and 95% confidence intervals for experimental values ($p = 0.99$, Fisher's exact test). **(f)** Relative frequency (expressed as percentage over all modeled translaminal lineages within a single simulation) of laminar configurations in experimental and modeled translaminal lineages. Grey boxes represent variability among 100 simulations; colored stars and lines show experimental values and 95% confidence intervals for experimental values ($p = 0.24$, Chi-square test). **(g)** 60 μm coronal sections through the primary somatosensory (S1) and visual (V1) cortex of P21 *Neurod6^{Cre/+};RCL-Fucci2* mice. The schemas on the right illustrate PC densities per layer. **(h)** Fraction of PCs per layer (expressed as percentage of total neurons) generated with two sets of laminar probability factors using Model 2 compared to the experimental data. Histograms represent mean \pm standard deviation. Z-scores represent the distance between experimental and simulated results for each parameter, which is calculated as the difference between the averages of model and experimental data divided by the standard deviation within model simulations (see methods for details). $n = 103$ neurogenic lineages in the somatosensory cortex of 25 animals. Data used for quantitative analyses as well as the numerical data that are represented in graphs are available in [Figure 4—source data 1](#), [Figure 7—source data 1](#) and [Figure 1—source data 2](#). The code for the generation of lineages can be found in [Figure 7—source data 2](#). See also [Figure 7—figure supplement 1](#).

Figure 7—source data 1. Laminar densities in *Neurod6^{Cre/+};RCL-Fucci2* mice analyzed at P21.

Figure 7–source data 2. Lineage generation simulator.

Figure 7–figure supplement 1. Stochastic models considering single and multiple programs corroborate Bayesian inference. **(a)** Schematic representation of a mathematical model of cortical neurogenesis in which a unique progenitor identity is modeled (Model 1). Squares represent the maximum number of stochastic decisions performed by each progenitor for each cortical layer during *in silico* simulations. The odds of generating neurons for each chance are given by a probability value (P), which is unique for each layer and progenitor identity. The model runs 100 simulations with 100 progenitors. **(b)** Fraction of cells in each cortical layer (expressed as percentage of total) in experimental and modeled lineages. **(c)** Clonal size distribution in experimental and modeled lineages. **(d)** Spearman correlation (r) values for the fraction of superficial and deep layer neurons in modeled lineages. Each dot represents an r value for one simulation. The green line shows the experimental value; the shadow area around the experimental data represents a 95% confidence interval for the experimental value. **(e)** Fraction of translaminal, deep and superficial layer-restricted lineages found experimentally and predicted by the model (expressed as percentage of all modeled lineages within a single simulation). Grey boxes represent variability among 100 simulations; colored stars and lines show experimental values and 95% confidence intervals for experimental values ($p = 0.13$, Fisher's exact test). **(f)** Relative frequency (expressed as percentage over all modeled translaminal lineages within a single simulation) of laminar configurations in experimental and modeled translaminal lineages. Grey boxes represent variability among 100 simulations; colored stars and lines show experimental values and 95% confidence intervals for experimental values ($p = 0.16$, Chi-square test). Histograms represent mean \pm standard deviation. Z-scores represent the distance between experimental and simulated results for each

1289 parameter, which is calculated as the difference between the averages of model and
1290 experimental data divided by the standard deviation within model simulations (see
1291 methods for details). $n = 103$ neurogenic lineages in the somatosensory cortex of 25
1292 animals. Data used for quantitative analyses as well as the numerical data that are
1293 represented in graphs are available in *Figure 4—source data 1* and *Figure 1—source data*
1294 *2*.

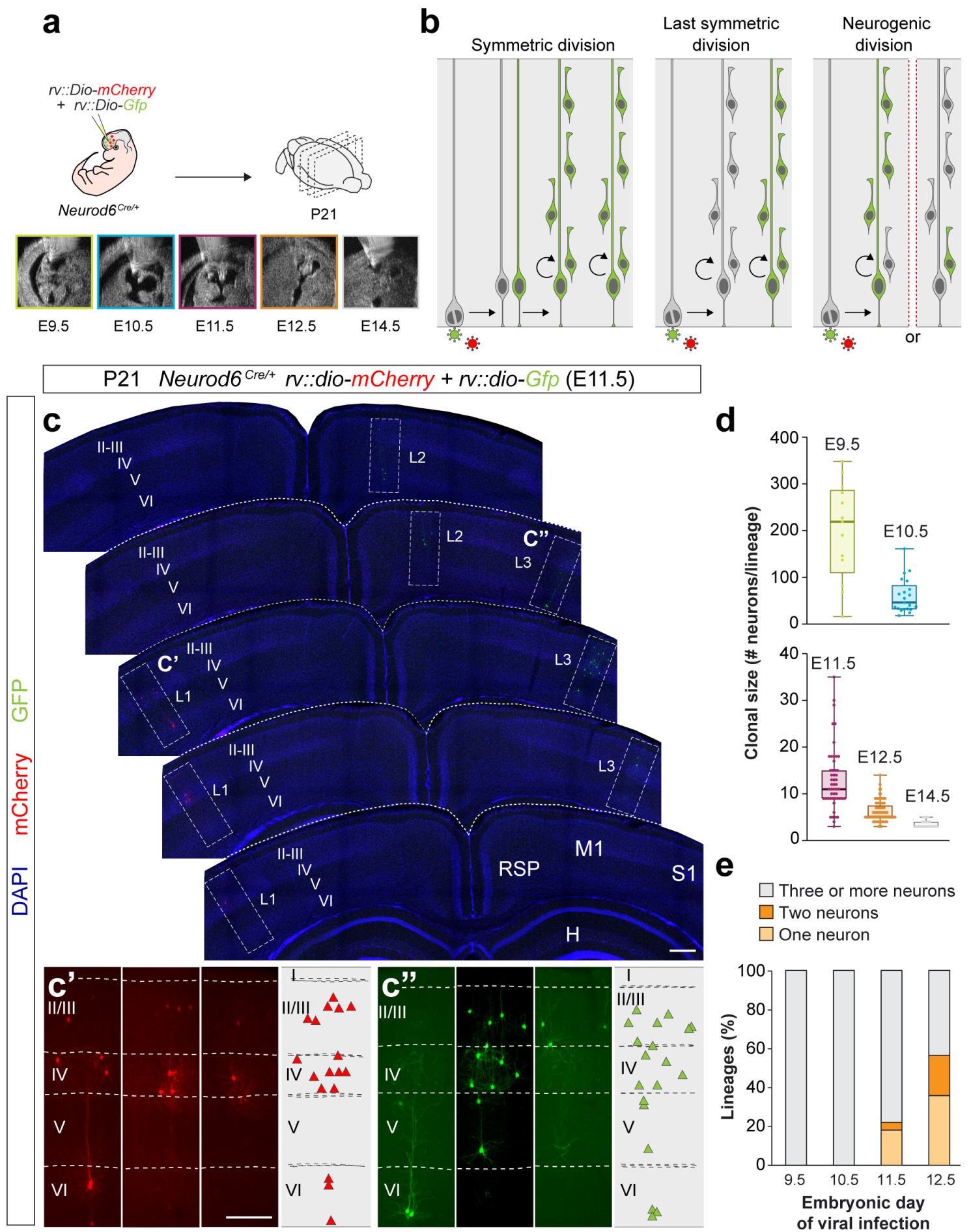


Figure 1

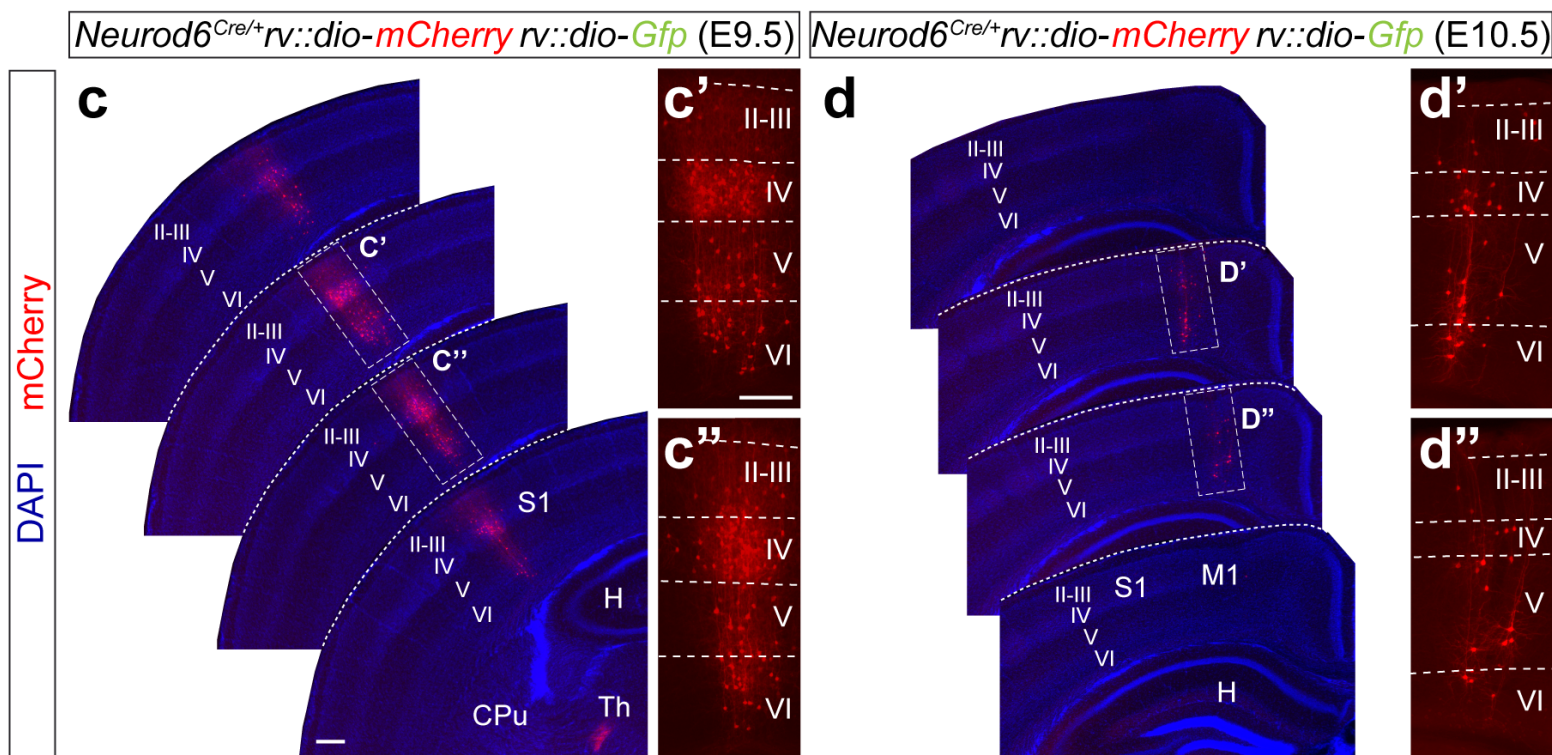
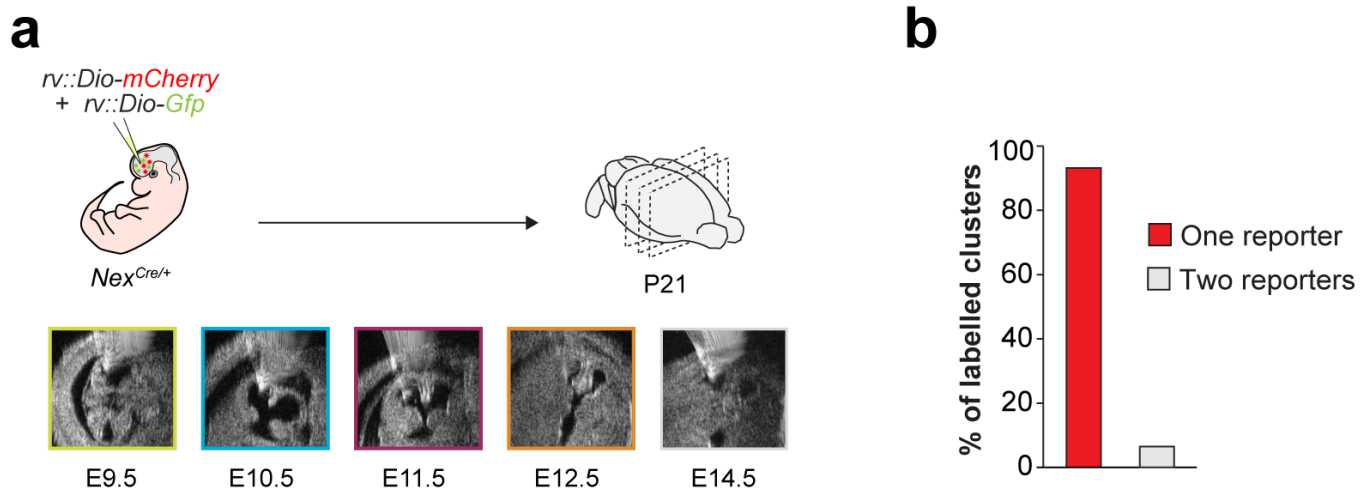


Figure 1-figure supplement 1

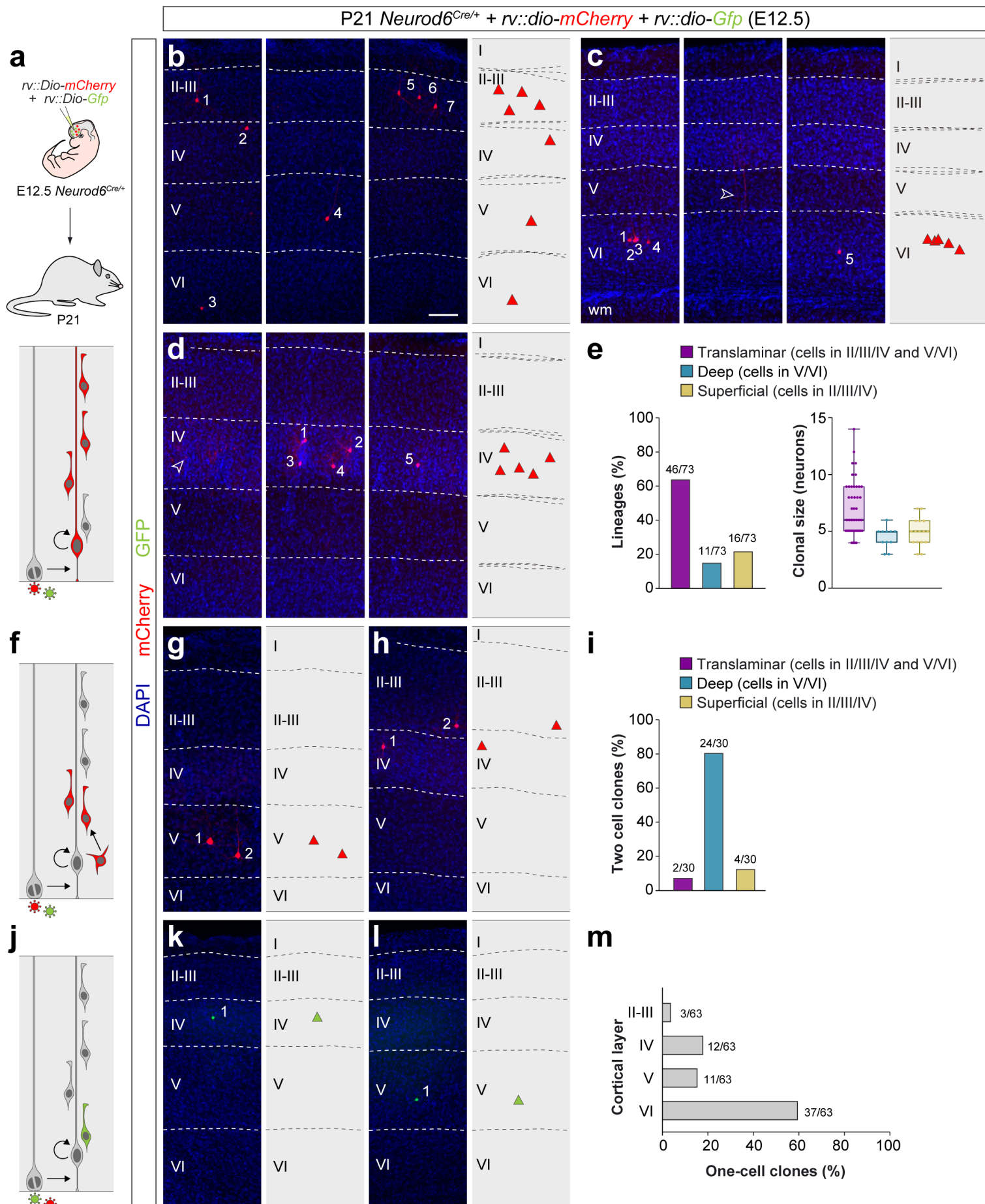


Figure 2

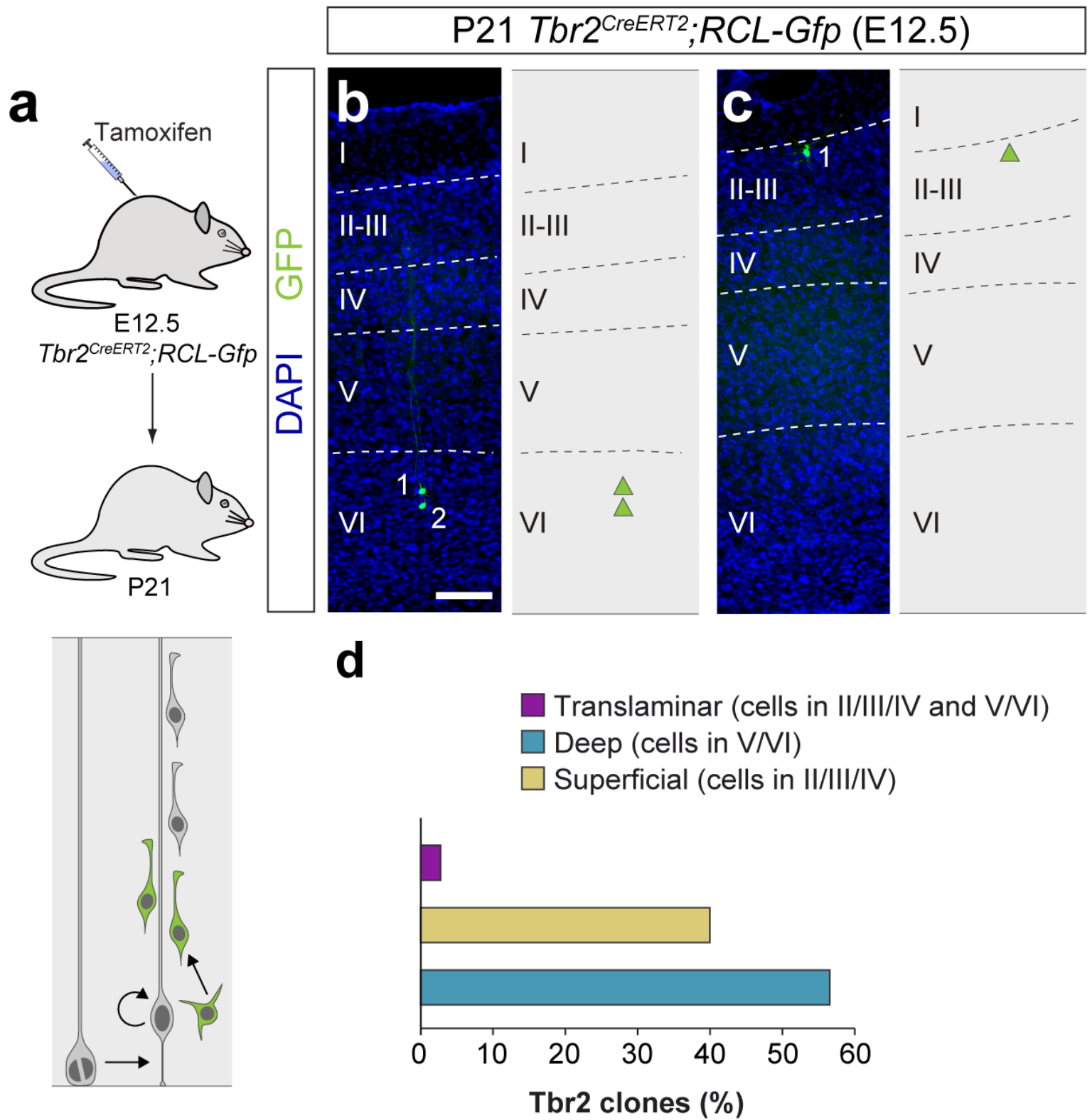


Figure 2-figure supplement 1

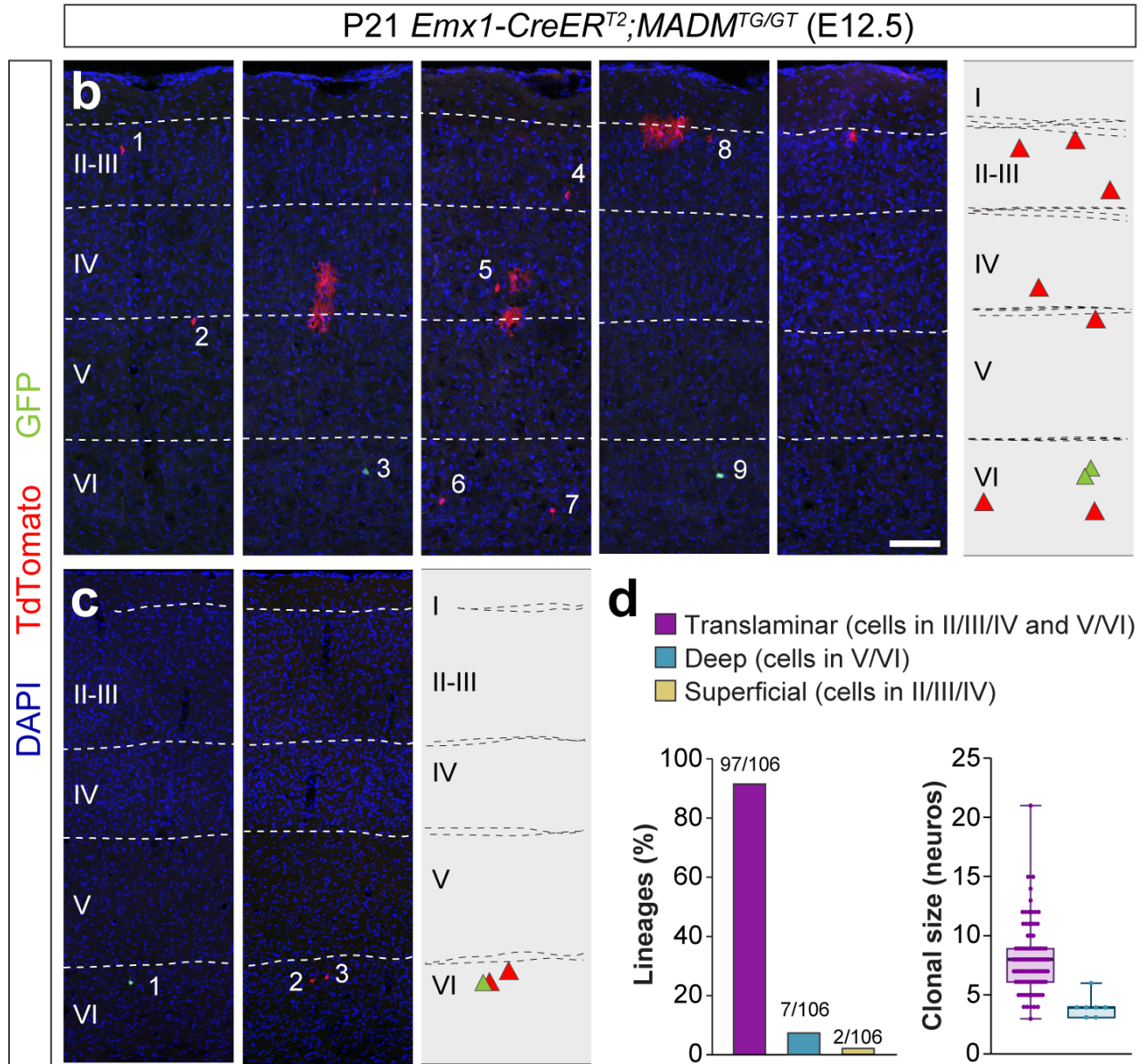
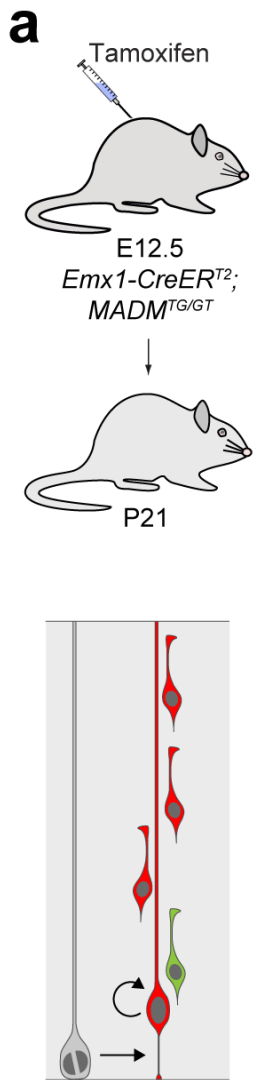
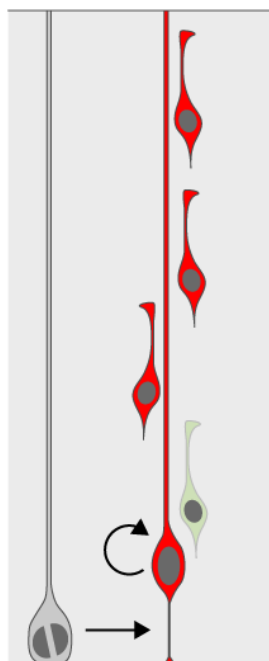


Figure 3



P21 *Emx1-CreER^{T2}*;MADM^{TG/GT} (E12.5)

b

Legend: GFP (green), tdTomato (red), DAPI (blue)

Layers: I, II-III, IV, V, VI

Cell numbers: 1, 2, 3, 4, 5, 6, 7

c

Legend: Translaminar (cells in II/III/IV and V/VI), Deep (cells in V/VI), Superficial (cells in II/III/IV)

Large subclones (%)

Category	Count	Percentage (%)
Translaminar (cells in II/III/IV and V/VI)	159/197	~80
Deep (cells in V/VI)	14/197	~7
Superficial (cells in II/III/IV)	24/197	~12

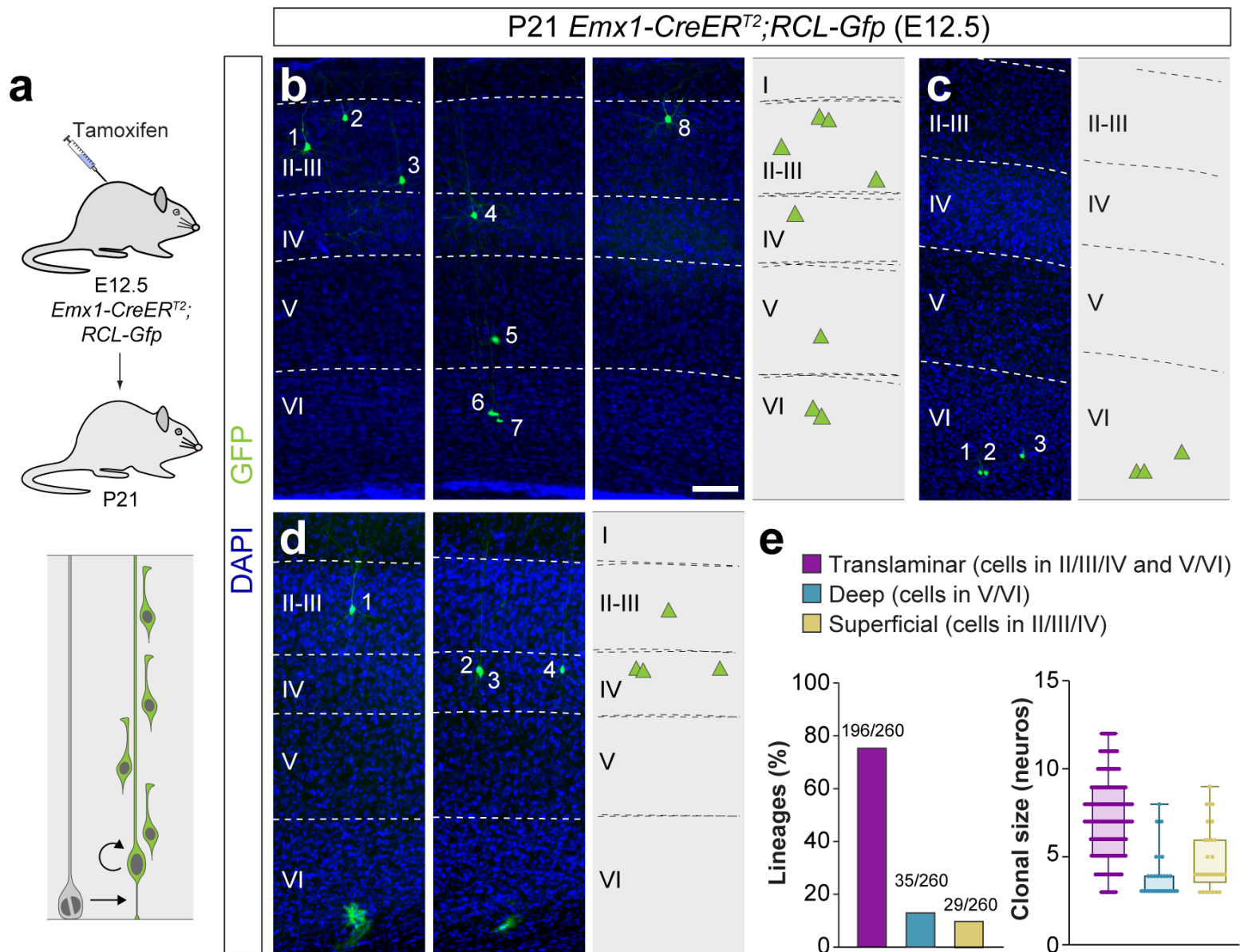


Figure 4

P21 *Emx1-CreER^{T2};RCL-Gfp* (E12.5)

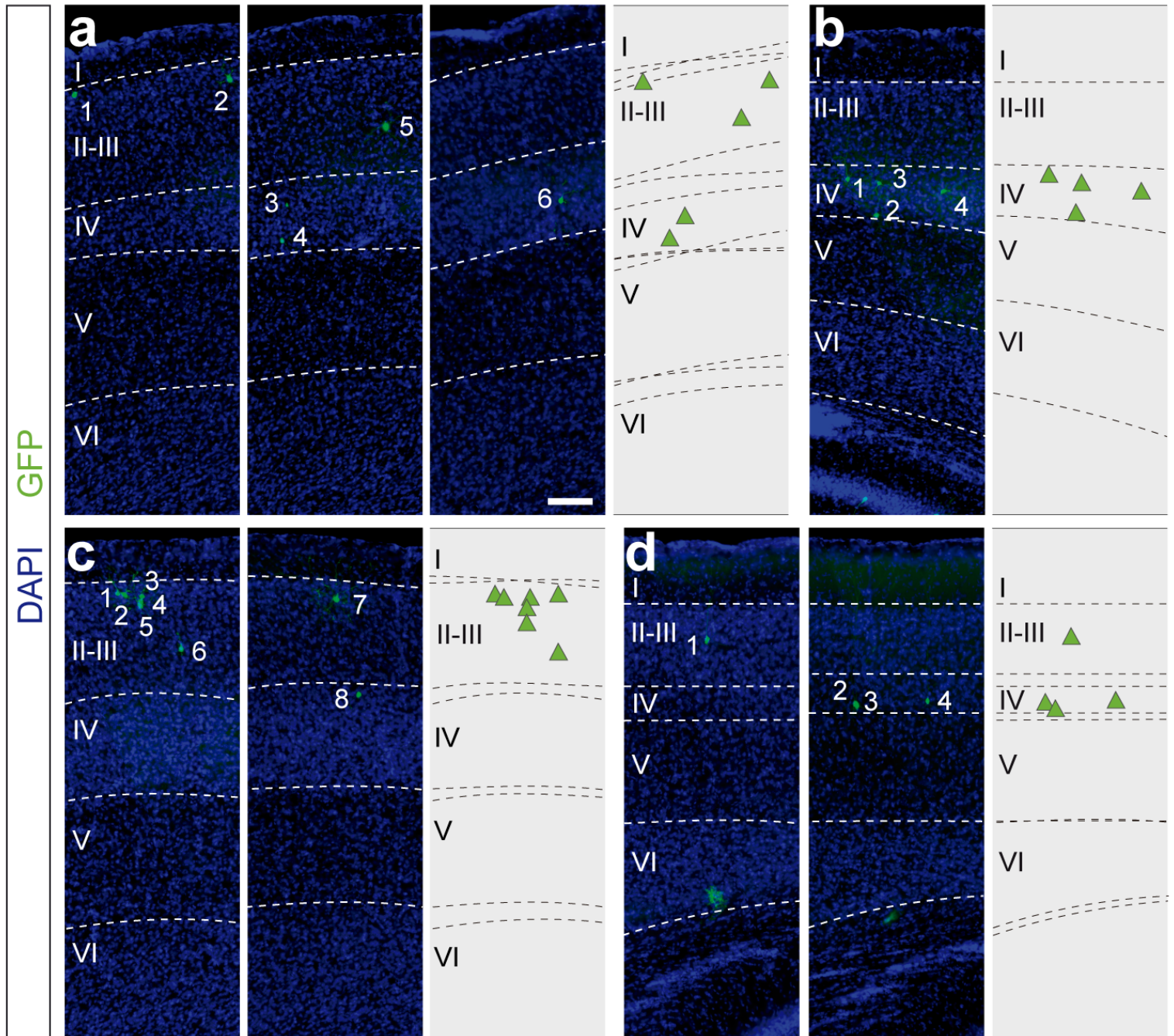


Figure 4-figure supplement 1

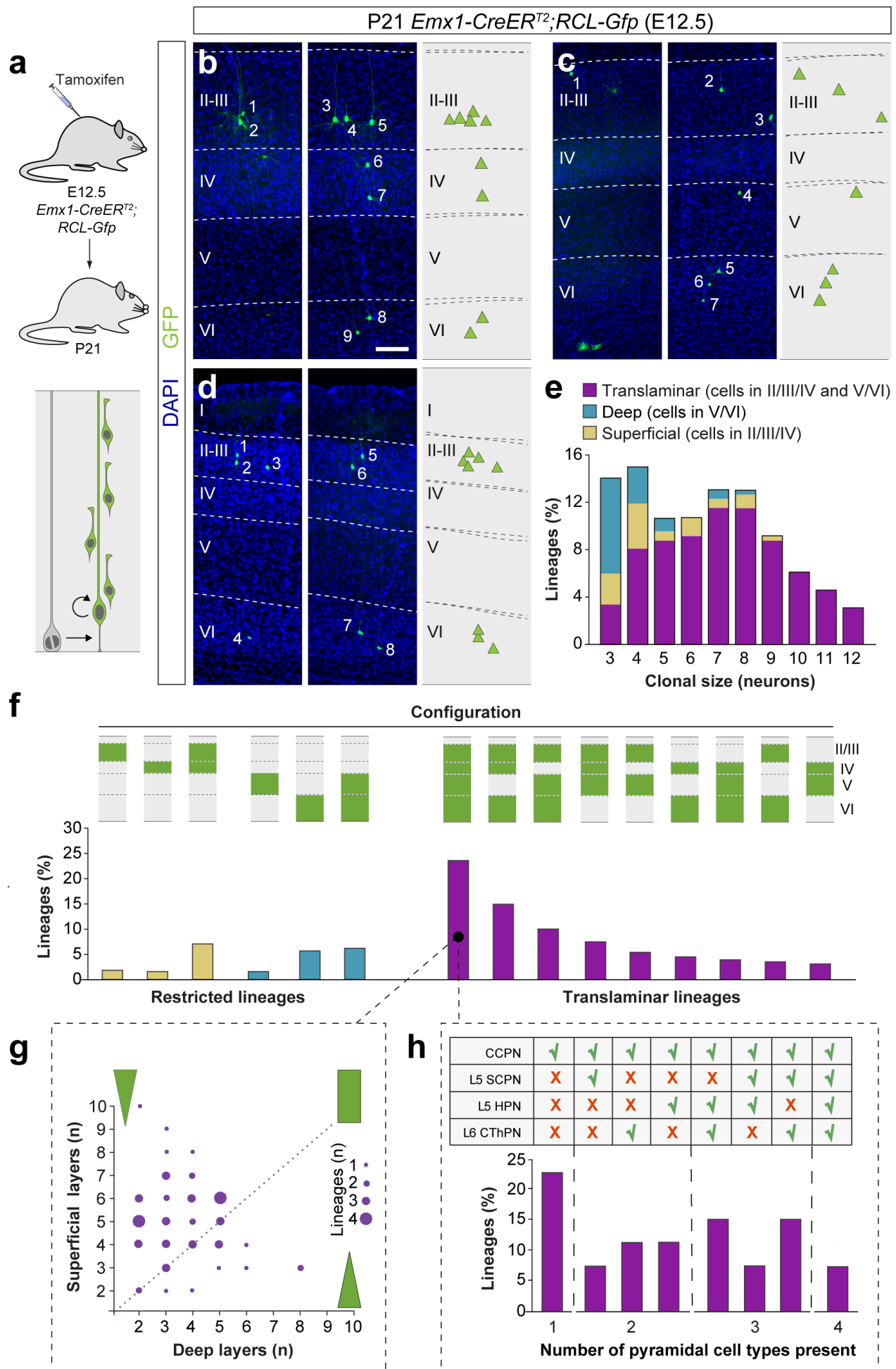


Figure 5

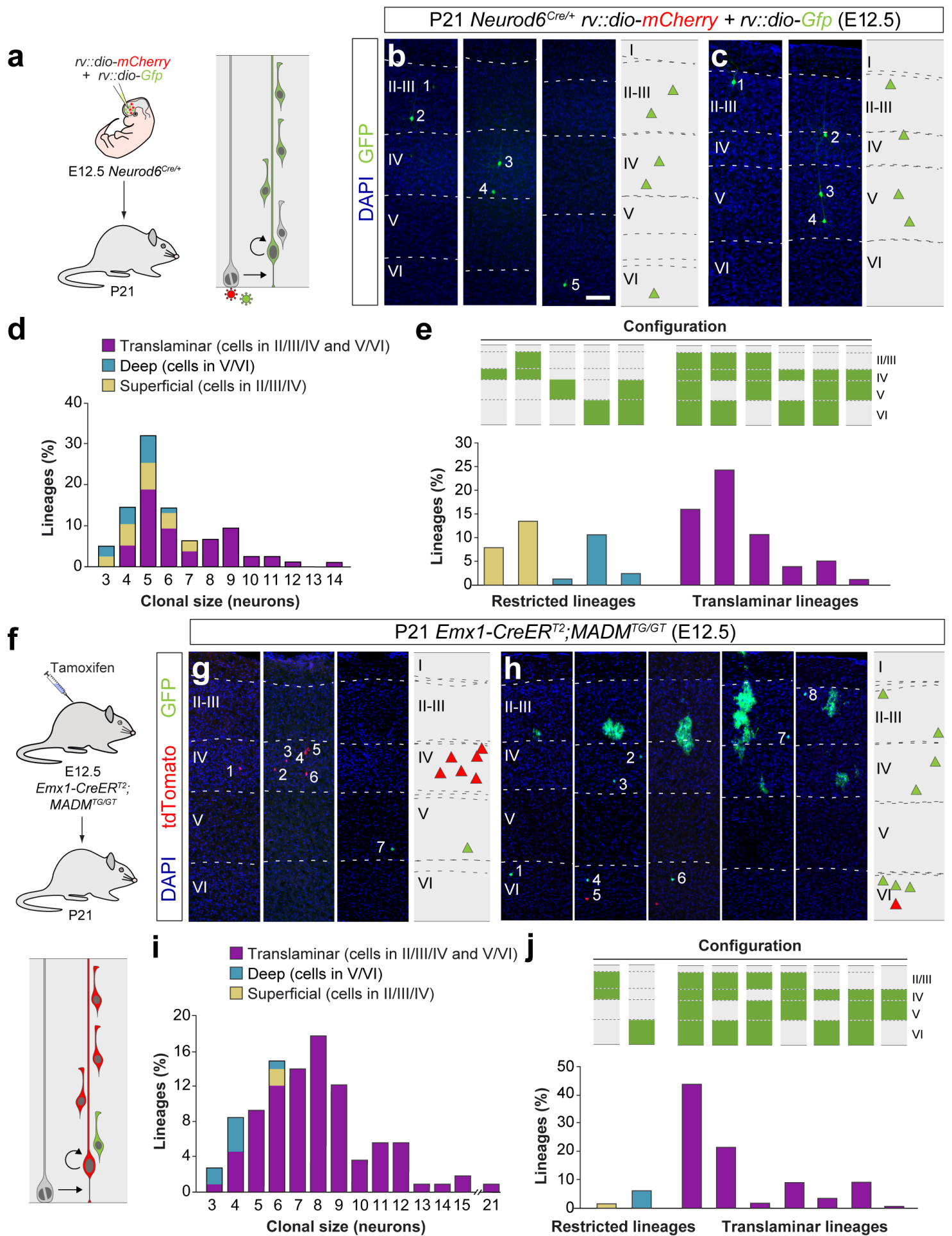


Figure 5-figure supplement 1

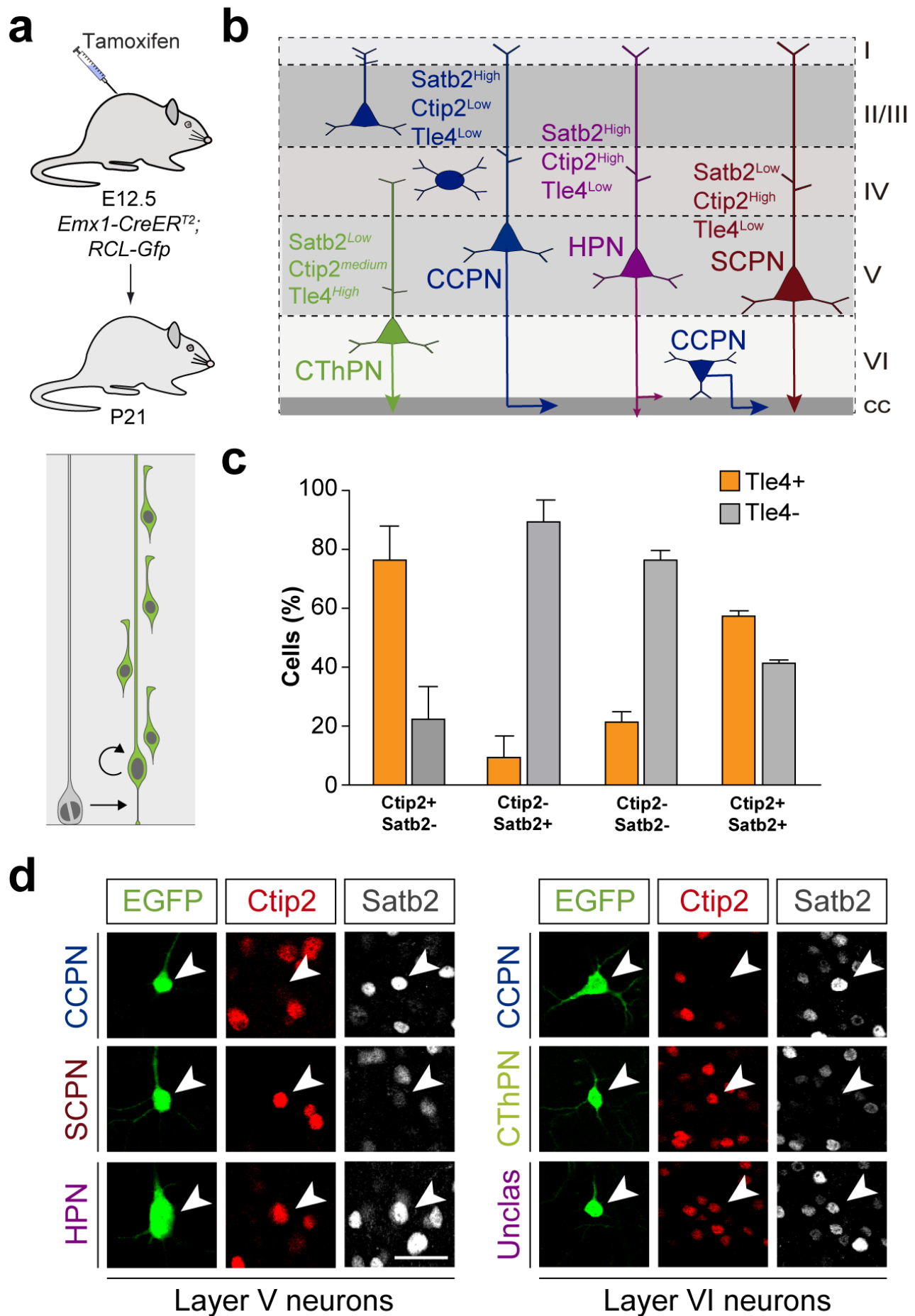


Figure 5-figure supplement 2

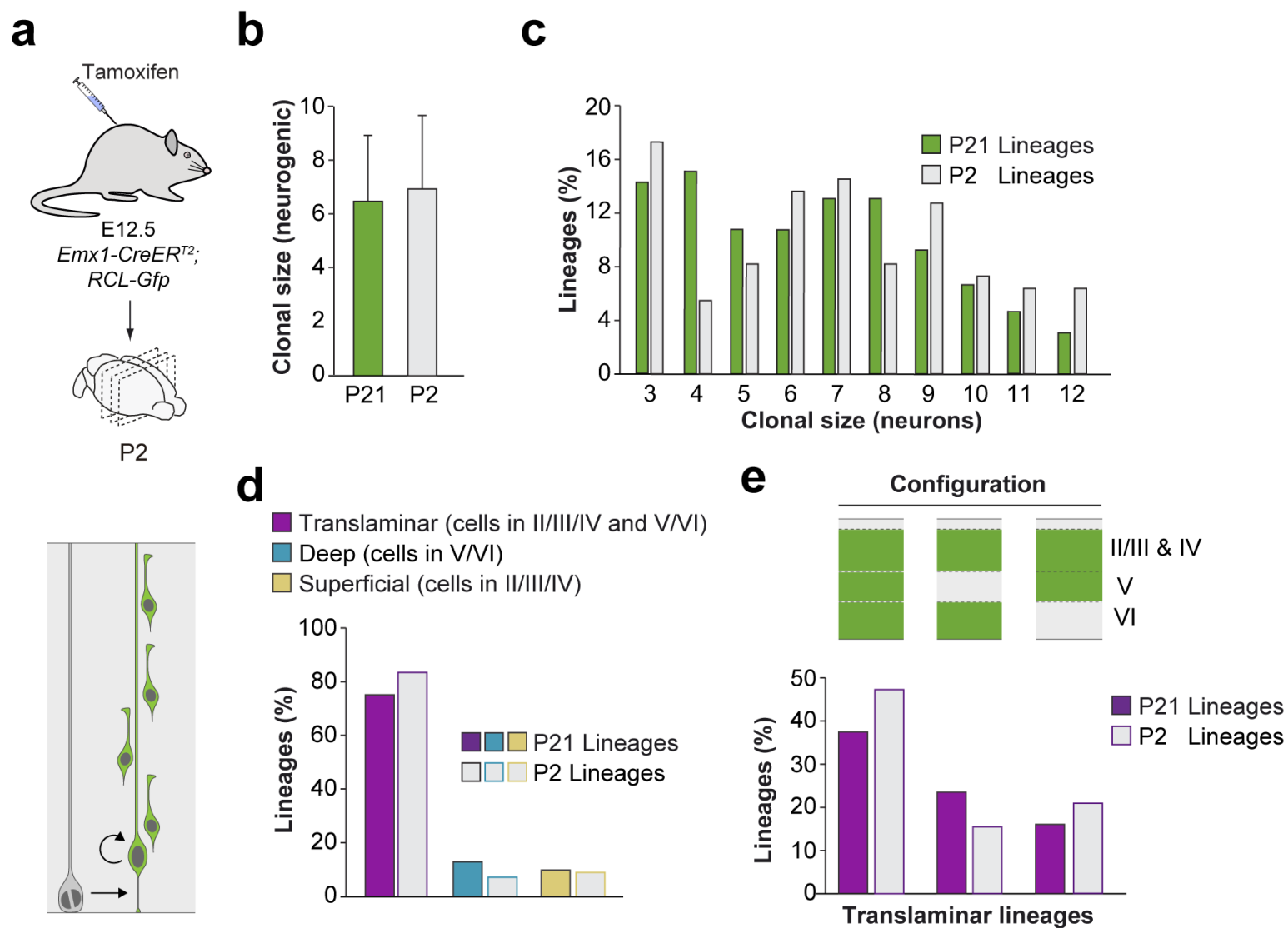


Figure 5-figure supplement 3

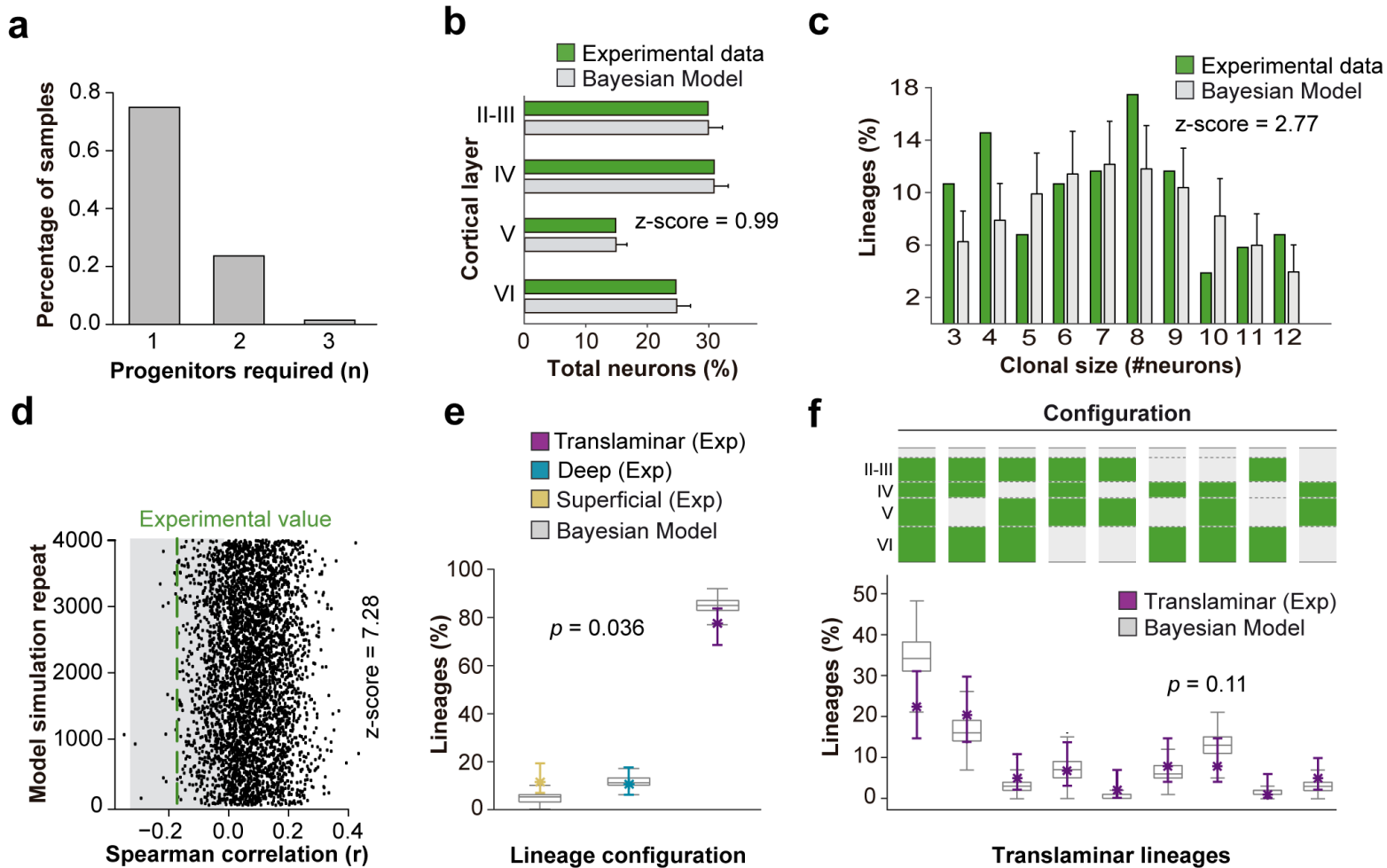


Figure 6

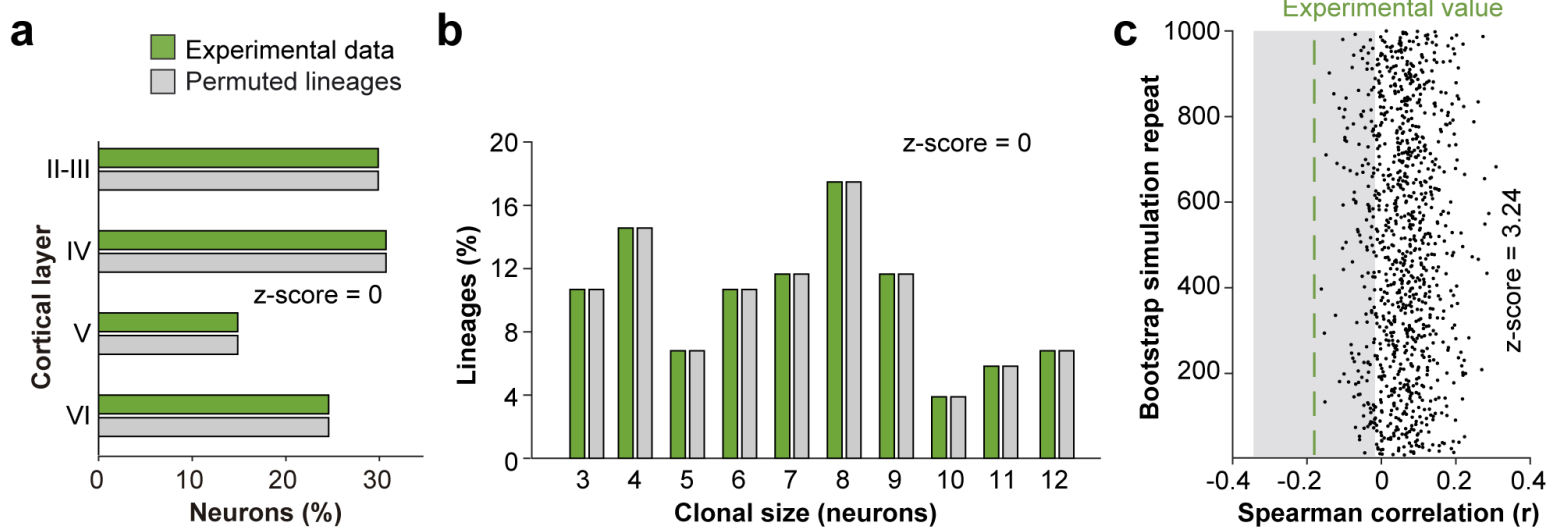


Figure 6-figure supplement 1

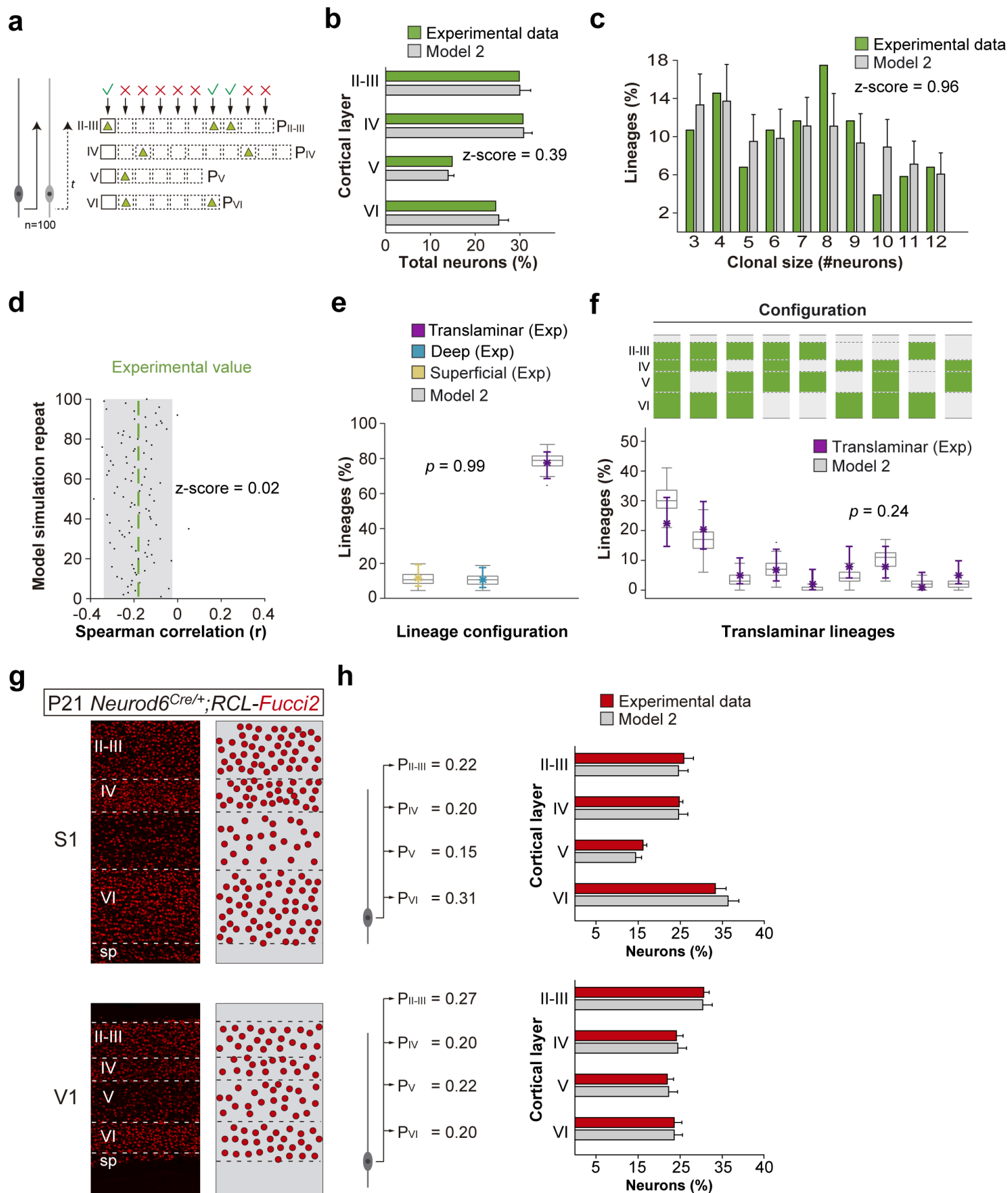


Figure 7

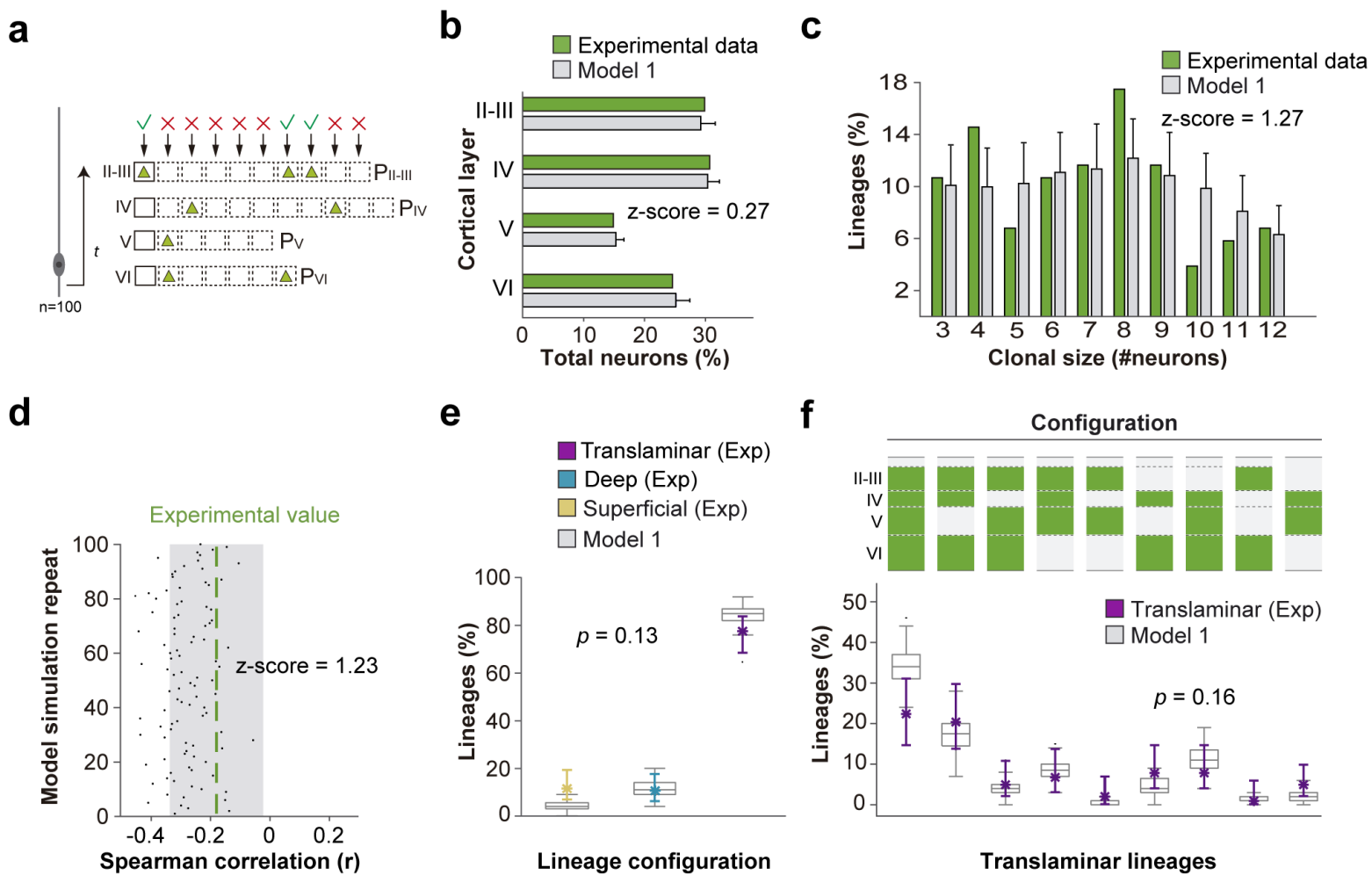


Figure 7-figure supplement 1